PROBABILITES

1 Introduction

1.1 Espace probabilisable et loi de variable aléatoire

1.1.1 Un exemple

1.1.2 Tribus

Définition 1.1.1 *On dit qu'une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ est une tribu si*

(a) $\Omega \in \mathcal{A}$,

(b) $A \in \mathcal{A} \implies \overline{A} \in \mathcal{A}$,

(c) $A_n \in \mathcal{A}, \forall n \in \mathbb{N} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Définition 1.1.2 *On appelle* **espace mesurable** *(ou mesurable) un couple* $(\Omega, \mathcal{A})$ *où* Ω *est un ensemble et* $\mathcal{A}$ *une tribu sur* Ω*. Les parties de* $\mathcal{A}$ *s'appellent les parties mesurables (ou les* **événements***) de l'espace mesurable* $(\Omega, \mathcal{A})$*.*

Théorème 1.1.3 *L'image réciproque d'une tribu par une application* f *est une tribu.*

Théorème 1.1.4 *Soit* $(\Omega, \mathcal{A})$ *un espace mesurable et* Ω' *une partie de* Ω*. L'ensemble* $\mathcal{A} \cap \Omega' = \{A \cap \Omega' \,/ A \in \mathcal{A}\}$ *est alors une tribu sur* Ω' *appelée* **tribu trace** *de* $\mathcal{A}$ *sur* Ω'*. Remarquons que la notation* $\mathcal{A} \cap \Omega'$ *est une convention.*

Théorème 1.1.5 *Soit* $(\mathcal{A}_i)_{i\in I}$ *une famille de tribus sur le même espace fondamental* Ω*. Alors* $\mathcal{A} = \bigcap\limits_{i\in I} \mathcal{A}_i$ *est une tribu sur* Ω*.*

Théorème 1.1.6 *Soit* $\mathcal{F}$ *une famille de parties de* Ω*. Il existe une plus petite tribu sur* Ω *qui contient* $\mathcal{F}$*. On l'appelle* **tribu engendrée** *par* $\mathcal{F}$ *et on la note* $\sigma(\mathcal{F})$*.*

Définition 1.1.7 *On appelle* **tribu borélienne** *sur* $\mathbb{R}$ *la tribu engendrée par les intervalles ouverts de la forme* $(-\infty, x)$*, pour tout* x *dans* $\mathbb{R}$*. On la note* $\mathcal{B}_{\mathbb{R}}$*.*

Théorème 1.1.8 *La tribu borélienne est également engendrée par les intervalles de la forme* $(-\infty, x]$*,* $(x, +\infty)$*,* $[x, +\infty)$*,* $[x, y]$*,* (x, y)*,* $[x, y)$*,* $(x, y]$*.*

1.1.3 Mesures et probabilités

Définition 1.1.9 *Soit* $(\Omega, \mathcal{A})$ *un espace mesurable. On appelle* **mesure positive** *sur* $(\Omega, \mathcal{A})$, *toute fonction* μ *définie sur* $(\Omega, \mathcal{A})$ *et à valeurs dans* $\overline{\mathbb{R}}^{+} = [0, +\infty]$, *telle que :*

(a) $\mu(\emptyset) = 0$,

(b) $\mu\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = \sum\limits_{n\in\mathbb{N}} \mu(A_n)$ *pour toute famille* $(A_n)_{n\in\mathbb{N}}$ *d'évènements disjoints deux à deux.*

Le triplet $(\Omega, \mathcal{A}, \mu)$ *est appelé* **espace mesuré**.

Définition 1.1.10 *On appelle* **probabilité** *sur* $(\Omega, \mathcal{A})$, *une mesure* P *sur* $(\Omega, \mathcal{A})$ *telle que* $P(\Omega) = 1$. *Le triplet* $(\Omega, \mathcal{A}, P)$ *est appelé* **espace probabilisé.**

Proposition 1.1.11 *Une probabilité vérifie les assertions suivantes :*

(a) $P\left(\overline{A}\right) = 1 - P(A)$, $\forall A \in \mathcal{A}$;

(b) Formule de Poincaré : si $(A_i)_{i\in\mathbb{N}} \subset \mathcal{A}$,

$$
\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} P(A_i) - \sum_{1\leq i_1<i_2\leq n}^{n} P\left(A_{i_1} \cap A_{i_2}\right) \\
&+ \sum_{1\leq i_1<i_2<i_3\leq n}^{n} P\left(A_{i_1} \cap A_{i_2} \cap A_{i_3}\right) \\
&+ \ldots + (-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n}^{n} P\left(A_{i_1} \cap \ldots \cap A_{i_3}\right) \\
&+ \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n);
\end{aligned}
$$

(c) Si A*,* $B \in \mathcal{A}$ *sont tels que* $A \subset B$*,*

$$P(A) \leq P(B);$$

(d) Inégalité de Boole : si $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$*,*

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

Définition 1.1.12 *On dit qu'une suite* $(A_n)_{n\in\mathbb{N}}$ *d'événements est croissante (resp. décroissante) si*

$$A_n \subset A_{n+1} \quad (resp.\ A_n \supset A_{n+1})$$

pour tout $n \in \mathbb{N}$.

Proposition 1.1.13 *Soit* P *une probabilité sur* $(\Omega, \mathcal{A})$.

(a) Si $(A_n)_{n\in\mathbb{N}^*} \subset \mathcal{A}$ *est une suite croissante d'événements, alors*

$$P\left(\bigcup_{n=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n) = \sup_{n\in\mathbb{N}^*} P(A_n).$$

(b) Si $(A_n)_{n\in\mathbb{N}^*} \subset \mathcal{A}$ *est une suite décroissante d'événements, alors*

$$P\left(\bigcap_{n=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n) = \inf_{n\in\mathbb{N}^*} P(A_n).$$

Définition 1.1.14 *Soit* $(\Omega, \mathcal{A}, \mu)$ *un espace mesuré tel que* $\{\omega\} \in \mathcal{A}$ *pour tout* $\omega \in \Omega$.

- *On dit que* μ *est une* **mesure discrète** *s'il existe une famille*

$$D = \{\omega_n : n \in I\}$$

(où I *est un ensemble d'indices fini ou dénombrable) d'éléments telle que*

$$\mu\left(\Omega \backslash D\right) = 0$$

et

$$\mu\left(A\right) = \mu\left(A \cap D\right) = \sum_{\omega_n \in A \cap D} \mu\left(\{\omega_n\}\right)$$

pour tout $A \in \mathcal{A}$.

- *On dit que* μ *est une* **mesure continue** *si elle ne possède pas d'atome, i.e. si pour tout* $\omega \in \Omega$ *on a* $\mu\left(\{\omega\}\right) = 0$.

Exemples

1. Mesure de Dirac
2. Probabilité sur Ω fini ou dénombrable
3. Mesure de comptage
4. Mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ et sur $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$

1.1.4 Variables aléatoires

Définition 1.1.15 *Soit* $(\Omega, \mathcal{A})$ *et* $(E, \mathcal{B})$ *deux espaces probabilisables. Une application* $f : \Omega \to E$ *est dite mesurable (ou* $(\mathcal{A}, \mathcal{B})$*-mesurable) si*

$$f^{-1}(\mathcal{B}) \subset \mathcal{A}.$$

Définition 1.1.17 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $(E, \mathcal{B})$ *un espace probabilisable. Une application mesurable* X *de* $(\Omega, \mathcal{A}, P)$ *vers* $(E, \mathcal{B})$ *est appelée* **variable aléatoire.**

Proposition 1.1.16

- *Si* f *et* g *sont deux fonctions mesurables de* $(\Omega, \mathcal{A})$ *vers* $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$*, alors les fonctions* $f+g$ *et* fg *sont encore mesurables.*

- *Si* f *et* g *sont deux fonctions mesurables de* $(\Omega, \mathcal{A})$ *vers* $(\Omega', \mathcal{A}')$ *et de* $(\Omega', \mathcal{A}')$ *vers* $(\Omega'', \mathcal{A}'')$ *respectivement, la fonction* $g \circ f$ *est mesurable de* $(\Omega, \mathcal{A})$ *vers* $(\Omega'', \mathcal{A}'')$*.*

- *Si* $(f_n)_n$ *est une suite de fonctions mesurables de* $(\Omega, \mathcal{A})$ *vers* $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$*, alors les fonctions*

$$\sup_n f_n, \ \inf_n f_n, \ \limsup_n f_n, \ \liminf_n f_n$$

sont mesurables, à condition qu'elles ne prennent pas de valeurs infinies.

1.1.5 Loi de probabilité d'une variable aléatoire

Définition 1.1.18 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* X *une variable aléatoire définie sur* $(\Omega, \mathcal{A})$ *et à valeurs dans* $(\Omega', \mathcal{A}')$*. On définit une fonction* P_X *sur* $\mathcal{A}'$ *et à valeurs dans* $[0, 1]$ *par :*

$$\forall A' \in \mathcal{A}' : P_X\left(A'\right) = P\left(X^{-1}\left(A'\right)\right).$$

P_X *est appelée* **loi de probabilité** *de la variable aléatoire* X*. On note également* $P_X(A') = P(X \in A')$ *pour tout événement* $A' \in \mathcal{A}'$*.*

1.2 Conditionnement

1.2.1 Probabilité conditionnelle à un événement

Définition 1.2.1 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. Soit* B *un évènement de* $\mathcal{A}$ *de probabilité non nulle :* $P(B) \neq 0$*. On appelle* **probabilité conditionnelle** *à* B*, la probabilité* P^B *sur* $(\Omega, \mathcal{A})$ *:*

$$
\begin{aligned}
P^B &: \mathcal{A} \to [0,1] \\
& A \mapsto P^B(A) = \frac{P(A \cap B)}{P(B)}.
\end{aligned}
$$

$P^B(A)$ *s'appelle probabilité conditionnelle de* A *à* B *ou encore probabilité conditionnelle de* A *sachant* B*. On note aussi*

$$P^B(A) = P(A/B).$$

Proposition 1.2.2
$$
\begin{aligned}
P(A \cap B) &= P(B)\, P(A/B) \quad \forall A, B \in \mathcal{A}. \\
&= P(A)\, P(B/A)
\end{aligned}
$$

1.2.2 Formule de Bayes

Soit $(A_i)_{i\in I}$ avec $I \subset \mathbb{N}$ une **partition** de Ω :

1. $\Omega = \bigcup\limits_{i\in I} A_i$
2. $\forall (i,j) \in I^2, i \neq j : A_i \cap A_j = \emptyset$

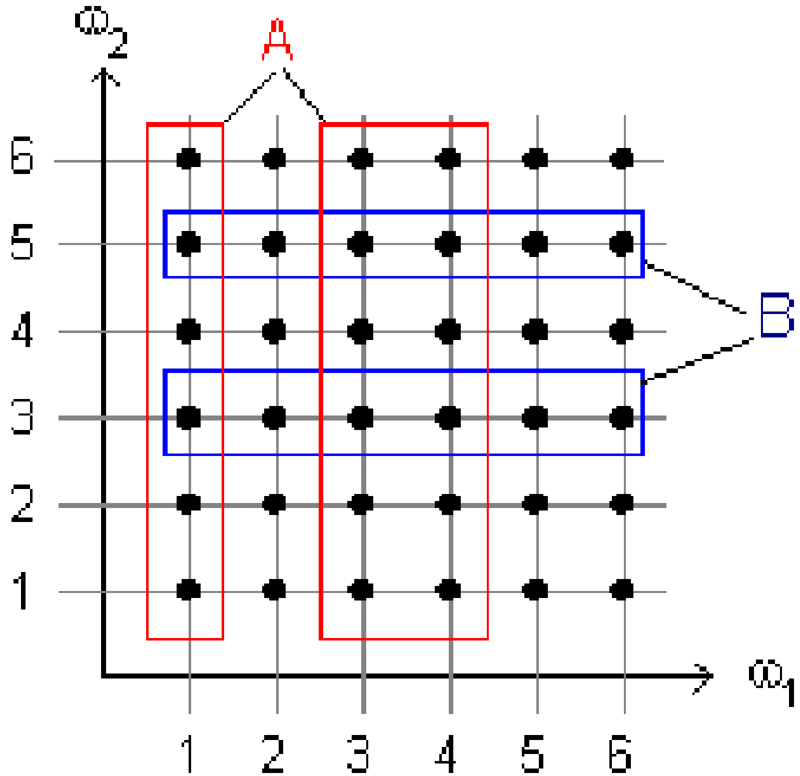
Supposons que :

$$\forall i \in I : P(A_i) \neq 0.$$

On a alors :

$$\forall j \in I : P\left(A_j / B\right) = \frac{P\left(A_j \cap B\right)}{P(B)} = \frac{P\left(A_j\right) P\left(B \middle/ A_j\right)}{\sum\limits_{i\in I} P(A_i)\, P(B / A_i)}.$$

Cette formule est appelée **formule de Bayes**.



ω_2
A
B
6
5
4
3
2
1
1
2
3
4
5
6
ω_1

1.3 Indépendance en probabilité

1.3.1 Indépendance d'événements

Définition 1.3.1 *On dit que deux évènements* A *et* B *d'un même espace probabilisé* $(\Omega, \mathcal{A}, P)$ *sont* **indépendants** *et on note* $A \perp\!\!\!\perp B$ *si :*

$$P(A \cap B) = P(A) P(B).$$

Proposition 1.3.2

(a) $A \perp\!\!\!\perp B \quad \Longrightarrow \quad \overline{A} \perp\!\!\!\perp B$, $A \perp\!\!\!\perp \overline{B}$ *et* $\overline{A} \perp\!\!\!\perp \overline{B}$.

(b) $P(A) \in \{0, 1\} \quad \Longrightarrow \quad A \perp\!\!\!\perp B$, $\forall B \in \mathcal{A}$

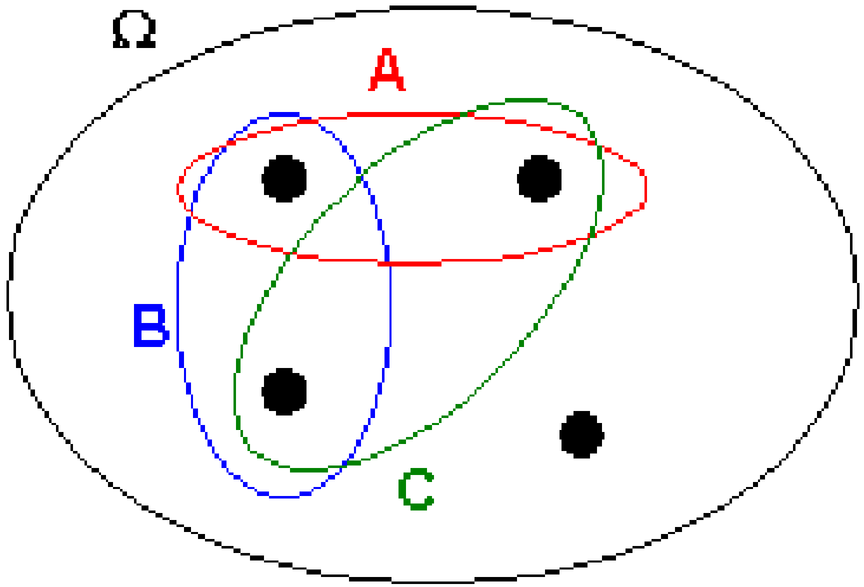
Définition 1.3.3 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. Soit* $(A_i)_{i=1,\dots,n}$ *une famille d'évènements de* $\mathcal{A}$*. Ces évènements sont dits* **(mutuellement) indépendants** *si :*

$$\forall J \subset \{1, \dots, n\} : P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P\left(A_j\right).$$

Remarque

L'indépendance mutuelle entraîne l'indépendance deux à deux de tous les évènements. La réciproque est fausse.

Définition 1.3.4 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. Une famille* $(A_i)_{i \in I} \subset \mathcal{A}$ *est une famille d'événements mutuellement indépendants si, pour tout ensemble d'indices* $K \subset I$ *fini, la famille* $(A_i)_{i \in K}$ *forme une famille d'événements mutuellement indépendants.*


Ω
A
B
C

1.3.2 Indépendance de tribus

Définition 1.3.5 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $(\mathcal{A}_i)_{i=1,\ldots,n}$ *une famille de sous-tribus de* $\mathcal{A}$. *La famille* $(\mathcal{A}_i)_{i=1,\ldots,n}$ *est dite indépendante si :*

$$P\left(A_1 \cap \ldots \cap A_n\right) = P\left(A_1\right) \times \ldots \times P\left(A_n\right), \quad \forall A_i \in \mathcal{A}_i, \quad \forall i \in \{1, \ldots, n\}.$$

1.3.3 Indépendance de variables aléatoires

Définition 1.3.6 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $\left(\Omega_i', \mathcal{A}_i'\right)_{i=1,\dots,n}$ *une famille d'espaces probabilisables. La famille de variables aléatoires* $(X_i)_{i=1,\dots,n}$*, où, pour tout* $i \in \{1, \dots, n\}$*,* $X_i : \Omega \longrightarrow \Omega_i'$*, est dite* **indépendante** *si :*

$$
\begin{aligned}
P\left(X_1 \in A_1', \dots, X_n \in A_n'\right) &= P\left(X_1 \in A_1'\right) \times \dots \times P\left(X_n \in A_n'\right), \\
\forall A_i' &\in \mathcal{A}_i', \quad \forall i \in \{1, \dots, n\}.
\end{aligned}
$$

Théorème 1.3.7 *Soit* $(\varphi_i)_{i=1,\dots,n}$ *une famille de fonctions mesurables respectivement définies sur* $\left(\Omega_i', \mathcal{A}_i'\right)_{i=1,\dots,n}$ *et à valeurs dans* $\left(\Omega_i'', \mathcal{A}_i''\right)_{i=1,\dots,n}$*. L'indépendance de la famille* $(X_i)_{i=1,\dots,n}$ *entraîne celle de la famille* $(\varphi_i(X_i))_{i=1,\dots,n}$*.*

Exemples

Soit X, Y, Z et T quatre v.a.r. indépendantes. On a alors :

1. $X^2 \perp\!\!\!\perp \exp(Y)$
2. $X + Y \perp\!\!\!\perp Z - T$

1.3.4 Lien entre les différents types d'indépendance

Proposition 1.3.8 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $(\mathcal{A}_i)_{i=1,\dots,n}$ *une famille de sous-tribus de* $\mathcal{A}$. *Il y a équivalence entre les assertions suivantes :*

(a) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} \mathcal{A}_i$

(b) $\underset{i\in I}{\perp\!\!\!\perp} X_i$ *pour toute famille* $(X_i)_{i=1,\dots,n}$ *de v.a avec* X_i $\mathcal{A}_i$*-mesurable*

(c) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} \mathbb{1}_{A_i}$ *pour toute famille* $(A_i)_{i=1,\dots,n}$ *d'évènements avec* $A_i \in \mathcal{A}_i$

(d) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} A_i$ *pour toute famille* $(A_i)_{i=1,\dots,n}$ *d'évènements avec* $A_i \in \mathcal{A}_i$.

Proposition 1.3.9 *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $(A_i)_{i=1,\ldots,n}$ une famille d'évènements sur $(\Omega, \mathcal{A}, P)$. Il y a équivalence entre les assertions suivantes :*

(a) $\underset{i=1,\ldots,n}{\perp\!\!\!\perp} A_i$

(b) $\underset{i=1,\ldots,n}{\perp\!\!\!\perp} \sigma(A_i)$

(c) $\underset{i=1,\ldots,n}{\perp\!\!\!\perp} \mathbb{1}_{A_i}$

Proposition 1.3.10 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. Soit* $\left(\Omega_i', \mathcal{A}_i'\right)_{i=1,\dots,n}$ *une famille d'espaces probabilisables. Soit* $(X_i)_{i=1,\dots,n}$ *une famille de v.a. respectivement définies sur* $(\Omega, \mathcal{A}, P)$ *et à valeurs dans* $\left(\Omega_i', \mathcal{A}_i'\right)_{i=1,\dots,n}$. *Il y a équivalence entre les assertions suivantes :*

(a) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} X_i$

(b) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} \sigma(X_i)$

(c) $\underset{i=1,\dots,n}{\perp\!\!\!\perp} \left\{X_i \in A_i'\right\}, \forall A_i' \in \mathcal{A}_i'$

Remarque

$\sigma(X_i) = X_i^{-1}\left(\mathcal{A}_i'\right)$ est la **tribu engendrée** par la v.a. X_i, c'est-à-dire la plus petite tribu rendant X_i mesurable.

1.4 Espace probabilisable produit

Définition 1.4.1 *Soit* $(\Omega_i, \mathcal{A}_i)_{i\in\{1,...,n\}}$ *une famille d'espaces mesurables. On appelle tribu produit des* $(\mathcal{A}_i)_{i\in\{1,...,n\}}$ *sur* $\Omega = \Omega_1 \times ... \times \Omega_n$ *et on note* $\bigotimes_{i=1}^{n} \mathcal{A}_i$ *la tribu engendrée par les pavés mesurables* $A_1 \times \ldots \times A_n$ *où* A_i *appartient à* $\mathcal{A}_i$ *pour tout* $i \in \{1, ..., n\}$ *:*

$$\bigotimes_{i=1}^{n} \mathcal{A}_i = \sigma\left(A_1 \times \ldots \times A_n : A_i \in \mathcal{A}_i,\ \forall i \in \{1, ..., n\}\right).$$

Théorème 1.4.2 *Soit* $(\Omega_i, \mathcal{A}_i, P_i)_{i\in\{1,...,n\}}$ *une famille d'espaces probabilisés. Il existe une probabilité unique* P *sur* $\left(\prod_{i=1}^{n}\Omega_i, \bigotimes_{i=1}^{n}\mathcal{A}_i\right)$ *telle que :*

$$P\left(A_1 \times ... \times A_n\right) = \prod_{i=1}^{n} P_i\left(A_i\right), \quad \forall A_i \in \mathcal{A}_i,\ \forall i \in \{1, ..., n\}.$$

Cette probabilité est appelée **probabilité produit** *des* P_i *et est notée* $\bigotimes_{i=1}^{n} P_i$.

Définition 1.4.3 *L'espace* $\left(\prod_{i=1}^{n}\Omega_i, \bigotimes_{i=1}^{n}\mathcal{A}_i, \bigotimes_{i=1}^{n}P_i\right)$ *est appelé* **espace probabilisé produit** *des espaces probabilisés* $(\Omega_i, \mathcal{A}_i, P_i)_{i\in\{1,...,n\}}$. *Si* $(\Omega_i, \mathcal{A}_i, P_i) = (\Omega, \mathcal{A}, P)$ *pour tout* $i \in \{1, ..., n\}$, *l'espace produit est noté*

$$(\Omega, \mathcal{A}, P)^{\otimes n}.$$

1.5 Loi conjointe d'un n-uplet de variables aléatoires indépendantes

Définition 1.5.1 *Soit $(X_i)_{i\in\{1,...,n\}}$ une famille de variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathcal{A},P)$ et respectivement à valeurs dans $(\Omega_i,\mathcal{A}_i)_{i\in\{1,...,n\}}$. On appelle* **loi conjointe** *du vecteur aléatoire $X = (X_1,...,X_n)$ la loi P_X de X sur $\left(\prod\limits_{i=1}^{n}\Omega_i, \bigotimes\limits_{i=1}^{n}\mathcal{A}_i\right)$. La loi P_{X_i} de chacune des variables aléatoires X_i est appelée* **loi marginale***.*

Proposition 1.5.2 *Les variables aléatoires $X_1,...,X_n$ sont indépendantes si et seulement si :*

$$P_X = \bigotimes_{i=1}^{n} P_{X_i}.$$

2 Variables et vecteurs aléatoires : lois sur $\mathbb{R}$ et sur $\mathbb{R}^n$

Rappels

- $(\Omega, \mathcal{A}, P)$ espace probabilisé
- $(\Omega', \mathcal{A}')$ espace mesurable
- $X : \Omega \to \Omega'$

X est une **variable aléatoire** (v.a.) :

$$X^{-1}\left(A'\right) \in \mathcal{A}, \quad \forall A^{'} \in \mathcal{A}'.$$

$(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$: X est une **variable aléatoire réelle** (v.a.r.)

Loi de probabilité de la v.a.r. X : la probabilité $P_X : \mathcal{B}_{\mathbb{R}} \to [0,1]$ définie par

$$P_X(I) = P(X \in I) = P\left(X^{-1}(I)\right) = P(\{\omega \in \Omega : X(\omega) \in I\})$$

pour tout $I \in \mathcal{B}(\mathbb{R})$.

2.1 Fonction de répartition

Définition 2.1.1 *On appelle* **fonction de répartition** *(f.r.) de la v.a.r.* X*, la fonction* F_X *définie sur* $\mathbb{R}$ *par :*

$$F_X(x) = P_X((-\infty, x]) = P(X \in (-\infty, x]) = P(X \leq x).$$

Proposition 2.1.2 *La f.r.* F_X *d'une v.a.r.* X *satisfait :*

(a) $0 \leq F_X(x) \leq 1, \quad \forall x \in \mathbb{R};$

(b) F_X *est croissante;*

(c) F_X *est continue à droite;*

(d) $\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to +\infty} F_X(x) = 1.$

Théorème 2.1.3 *Toute fonction F définie sur $\mathbb{R}$ qui vérifie les propriétés de la proposition précédente est une fonction de répartition d'une v.a.r.*

Proposition 2.1.4 *Le saut $p_0 = F_X(x_0) - F_X(x_0-)$ de la fonction de répartition F_X au point x_0 est égal à $P(X = x_0)$.*

Définition 2.1.5 *On appelle* **quantile** *d'ordre α (pour $0 < \alpha < 1$) de la loi X, tout réel x_α tel que :*

$$P\left(X \leq x_\alpha\right) \geq \alpha \quad et \quad P\left(X \geq x_\alpha\right) \geq 1 - \alpha.$$

Terminologie

Ordre des quantiles	**Nom des quantiles**
$1/2$	**Médiane** de X : $Med(X)$
$k/4,\ k \in \{1, 2, 3\}$	**Quartiles** : Q_1, Q_2, Q_3
$k/10, k \in \{1, ..., 9\}$	**Déciles**
$k/100, k \in \{1, ..., 99\}$	**Centiles**

- Si la médiane n'est pas unique alors on parle dans ce cas là d'intervalle médian.

- $[Q_1, Q_3]$ est appelé **espace interquartile**.

2.1.1 Lois discrètes

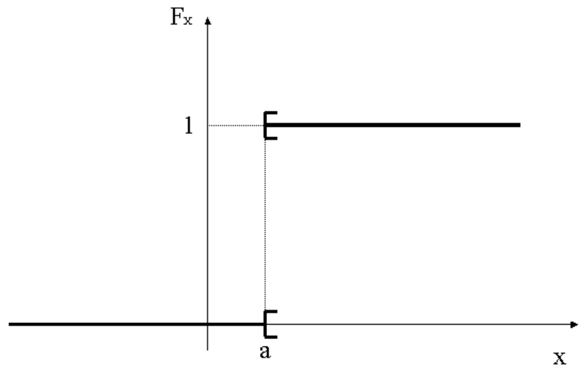
Définition 2.1.6 *On dit qu'une v.a.r.* X *est* **discrète** *si sa loi de probabilité* P_X *est une mesure de probabilité discrète.*

Proposition 2.1.7 *La fonction de répartition* F_X *d'une v.a.r. discrète* X *est une fonction en escalier dont les sauts sont situés sur les atomes.*

Loi de Dirac $X \sim \delta_a$, $a \in \mathbb{R}$: $P_X = \delta_a$,

$$P_X(A) = \delta_a(A) = \mathbb{1}_A(a), \quad \forall A \in \mathcal{B}_{\mathbb{R}}.$$

La fonction de répartition de la v.a.r. X a l'allure suivante :



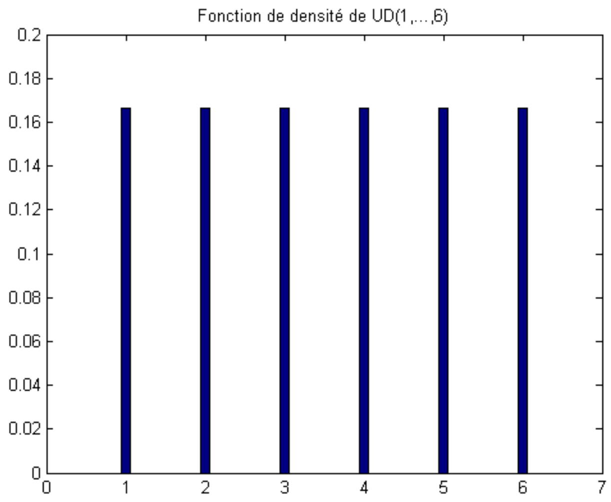
Loi uniforme discrète $X \sim \mathcal{U}_{\{x_1,\ldots,x_n\}}$:

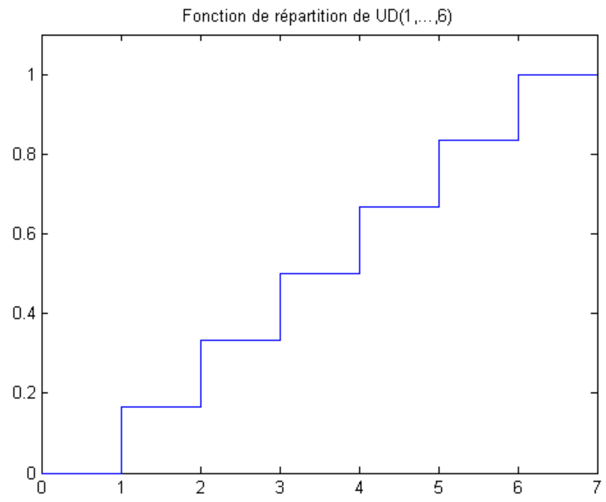
$$P_X\left(\{x_i\}\right) = P\left(X = x_i\right) = \frac{1}{n}, \quad \forall i \in \{1, \ldots, n\}.$$

Exemple

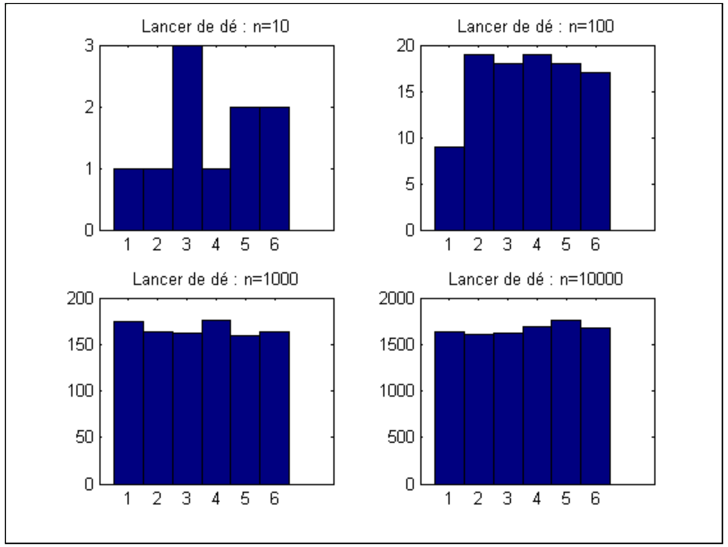
Soit le lancer d'un dé numéroté de 1 à 6. On peut considérer que la variable aléatoire X désignant le résultat du lancer de dé, suit une loi uniforme discrète sur $\{1, ..., 6\}$.

Ci-dessous figurent les fonctions de densité et de répartition de cette loi.





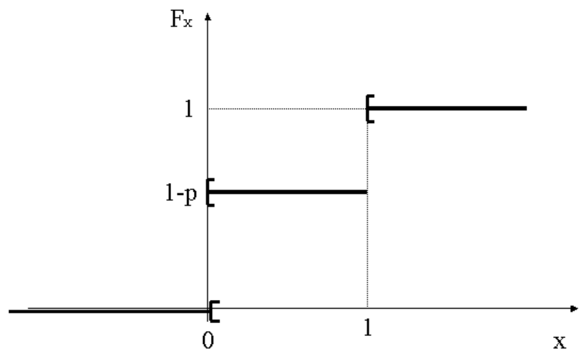
Afin de faire le lien entre modèle probabiliste et modèle statistique, figurent ci-après le résultat de simulations de lancers de dés. Il est facile de constater que plus le nombre de lancers augmente, plus la distribution empirique se rapproche de la distribution uniforme discrète, c'est à dire la loi théorique.



Loi de Bernoulli $X \sim \mathcal{B}(1,p)\,,\ p \in [0,1]$:

$$P_X(\{1\}) = P(X=1) = p$$
$$P_X(\{0\}) = P(X=0) = 1-p = q$$

La fonction de répartition de la v.a.r. X a l'allure suivante :



Soit le tirage d'une boule dans une urne contenant des boules blanches et des boules noires, en proportions respectivement égales à p et $q = 1 - p$. La v.a.r. X défini par :

$$X(\omega) = \begin{cases} 1 & si \ on \ tire \ une \ boule \ blanche \\ \\ 0 & si \ on \ tire \ une \ boule \ noire \end{cases}$$

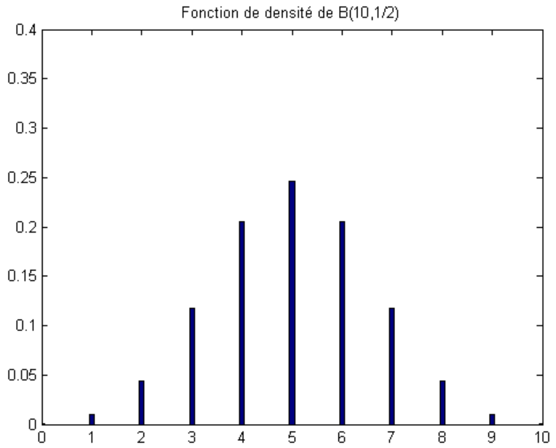
suit une **loi de Bernoulli** de paramètre $p \in [0, 1]$.

Loi binomiale $X \sim \mathcal{B}(n,p)\,,\ n \in \mathbb{N}^*,\ p \in [0,1]$:

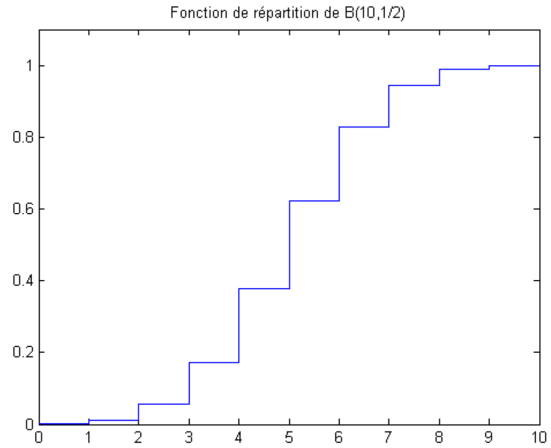
$$P_X(\{k\}) = P(X=k) = C_n^k p^k (1-p)^{n-k}\,, \quad \forall k \in \{0,...,n\}$$

Considérons n tirages successifs avec remise d'une boule dans une urne contenant des boules blanches et des boules noires, en proportions respectivement égales à p et $q = 1-p$. La v.a.r. X désignant le nombre de boules blanches tirées à l'issue de ces n tirages suit une **loi binomiale** de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$.

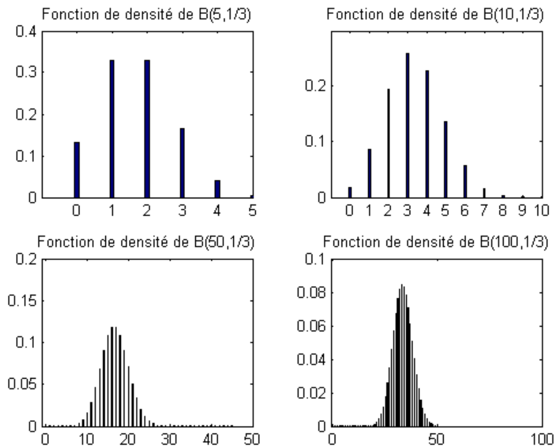
Exemples



Fonction de densité de B(10,1/2)

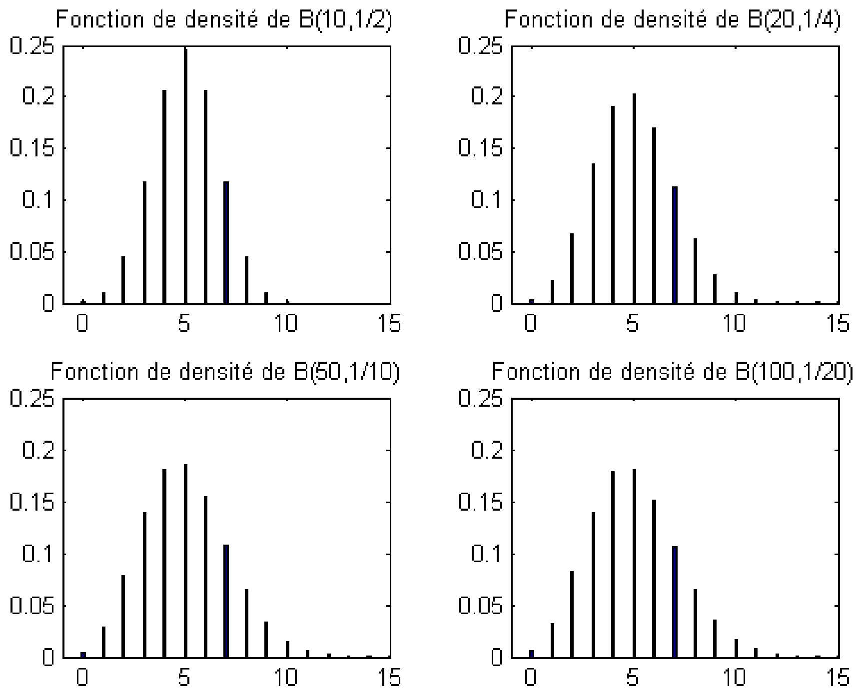


Fonction de répartition de B(10,1/2)



Fonction de densité de B(5,1/3)
Fonction de densité de B(10,1/3)
Fonction de densité de B(50,1/3)
Fonction de densité de B(100,1/3)

Voici les distributions pour différentes lois binomiales telles que leurs paramètres vérifie $np = 5$.



Remarques

1. Si :

$$X_1 \sim \mathcal{B}(n_1, p), \quad X_2 \sim \mathcal{B}(n_2, p), \quad X_1 \perp\!\!\!\perp X_2,$$

alors on a :

$$X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p).$$

En particulier, si $(X_i)_{i \in \{1,\ldots,n\}}$ est une famille de v.a.r. indépendantes, suivant une loi de Bernoulli $\mathcal{B}(1, p)$ de paramètre p, alors :

$$\sum_{i=1}^{n} X_i \sim \mathcal{B}(n, p).$$

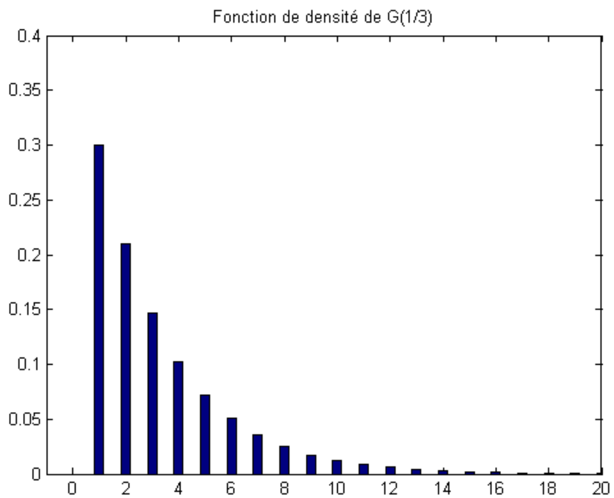
2. Si $X \sim \mathcal{B}(n, p)$ avec $p \in \left(\frac{1}{2}, 1\right]$ alors $n - X \sim \mathcal{B}(n, 1-p)$.

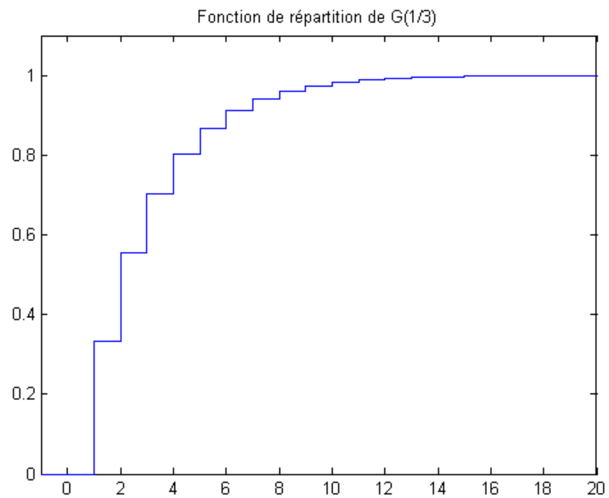
Ce résultat est utile pour lire les abaques des lois binomiales, souvent donnés pour $p \in \left[0, \frac{1}{2}\right]$.

Loi géométrique $\quad X \sim \mathcal{G}(p)\,,\ p \in [0,1]$:

$$P_X(\{k\}) = P(X = k) = (1-p)^{k-1}\,p, \quad \forall k \in \mathbb{N}^*$$

Soit une urne contenant des boules blanches et des boules noires, en proportions respectivement égales à p et $q = 1 - p$. La v.a.r. X désignant le nombre de tirages successifs avec remise effectués pour obtenir une boule blanche suit une **loi géométrique** de paramètre $p \in [0,1]$.





Loi binomiale négative $X \sim \mathcal{BN}(n,p)\,,\ n \in \mathbb{N},\ p \in [0,1]$:

$$P_X(\{k\}) = P(X = k) = C_{n+k-1}^{n-1} p^n (1-p)^k\,, \quad \forall k \in \mathbb{N}$$

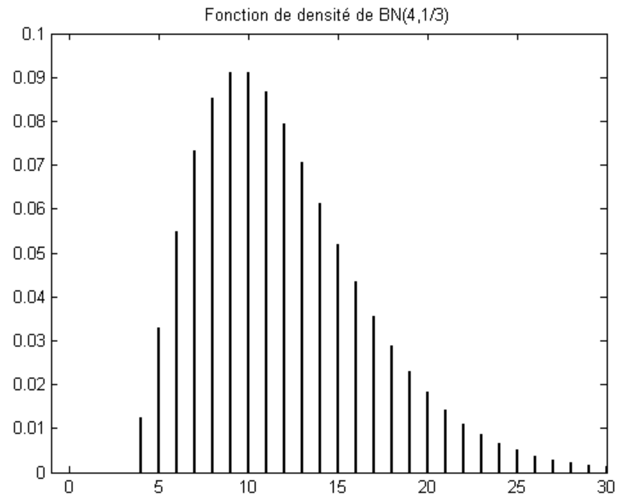
Soit une urne contenant des boules blanches et des boules noires, en proportions respectivement égales à p et $q = 1 - p$. Soit Y le nombre de tirages que l'on doit faire pour obtenir n boules blanches. La v.a.r. $X = Y - n$ désignant le nombre de boules noires obtenues avant d'avoir n boules blanches suit une **loi binomiale négative** de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$.

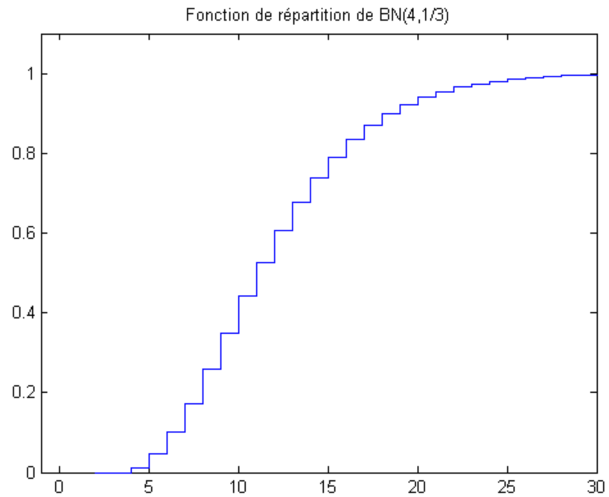
Remarque

Si $(X_i)_{i \in \{1,\dots,n\}}$ est une famille de v.a.r. indépendantes, suivant une loi géométrique $\mathcal{G}(p)$ de paramètre p alors :

$$\sum_{i=1}^{n} X_i \sim \mathcal{BN}(n,p)\,.$$

Exemple


Fonction de densité de BN(4,1/3)
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
20
25
30


Fonction de répartition de BN(4,1/3)
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30

Loi de Poisson $X \sim \mathcal{P}(\lambda)\,,\ \lambda > 0$:

$$P_X(\{k\}) = P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad \forall k \in \mathbb{N}.$$

Cette loi est souvent utilisée pour décrire des évènements rares comme le nombre d'accidents, d'erreurs de fabrication, d'individus atteints d'une maladie...

Remarque

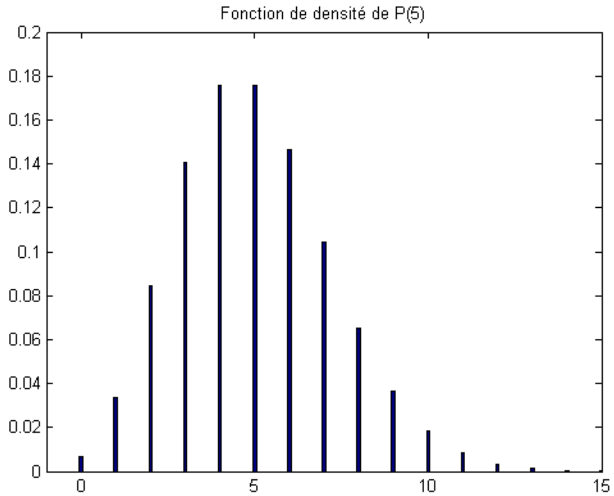
Si :

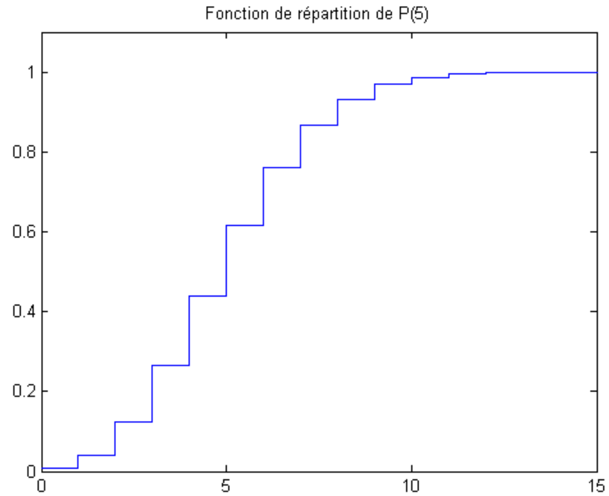
$$X_1 \sim \mathcal{P}(\lambda_1)\,, \quad X_2 \sim \mathcal{P}(\lambda_2)\,, \quad X_1 \perp\!\!\!\perp X_2,$$

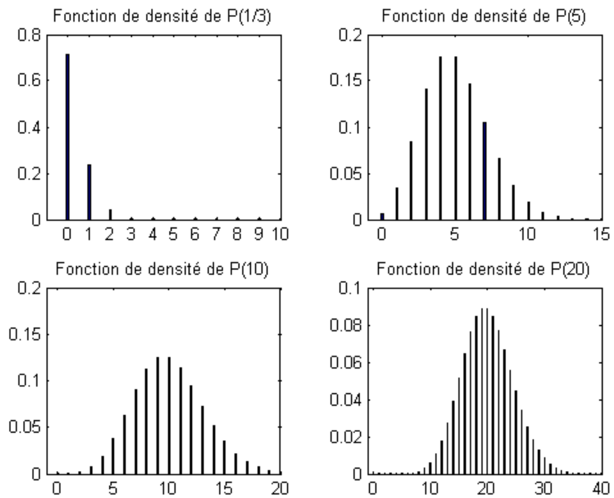
alors on a :

$$X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)\,.$$

Exemples


Fonction de densité de P(5)

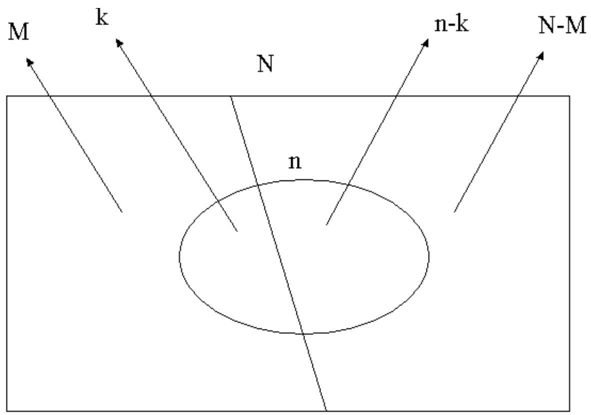

Fonction de répartition de P(5)


Fonction de densité de P(1/3)
Fonction de densité de P(5)
Fonction de densité de P(10)
Fonction de densité de P(20)

Loi hypergéométrique $\quad X \sim \mathcal{H}(n, N, M)\,,\; n, N, M \in \mathbb{N}^*\;$:

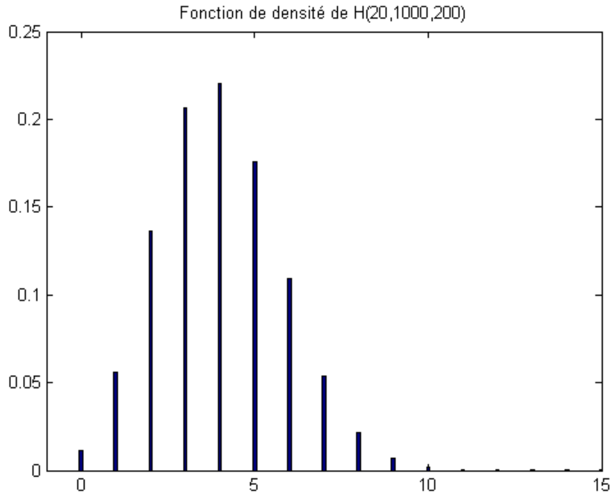
$$P_X(\{k\}) = P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad \forall k \in \{\max(0, n + M - N), ..., \min(n, M)\}$$

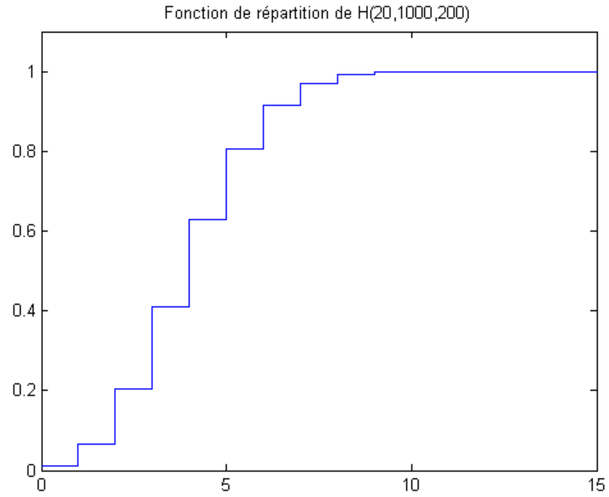
Soient n tirages successifs sans remise d'une boule dans une urne contenant N boules, dont M boules blanches et $N - M$ boules noires. La v.a.r. X désignant le nombre de boules blanches tirées à l'issue de ces n tirages suit une **loi hypergéométrique** de paramètres n, N et $M \in \mathbb{N}^*$.



.

Exemple


Fonction de densité de H(20,1000,200)
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15


Fonction de répartition de H(20,1000,200)
1
0.8
0.6
0.4
0.2
0
0
5
10
15

2.1.2 Lois continues

Définition 2.1.8 *On dit qu'une v.a.r.* X *est (de loi)* **continue** *si sa loi* P_X *est une mesure de probabilité continue.*

Proposition 2.1.9 *Une v.a.r. est continue si et seulement si sa fonction de répartition est continue.*

Définition 2.1.10 *On dit que la v.a.r.* X *est* **absolument continue** *si elle admet une densité de probabilité : il existe une fonction positive* f_X *telle que*

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$$

pour tout $x \in \mathbb{R}$*.*

Théorème 2.1.11 *Une fonction f définie sur $\mathbb{R}$ est une densité de probabilité si et seulement si elle vérifie :*

(a) f *positive,*

(b) f *mesurable,*

(c) f *intégrable et*

$$\int_{-\infty}^{+\infty} f(x)\ dx = 1.$$

Proposition 2.1.12 *Si la densité de probabilité f_X est continue sur un intervalle $[a, b]$ alors la fonction de répartition F_X est dérivable sur $[a, b]$ et on a $f_X = F_X^{'}$.*

Proposition 2.1.13 *Une v.a.r. absolument continue est continue mais la réciproque est fausse.*

Théorème 2.1.14 *Soit F une fonction de répartition. Alors il existe trois fonctions de répartitons F_1 discrète, F_2 absolument continue et F_3 singulière (i.e. continue mais pas absolument continue) et trois nombres réels α_1, α_2 et α_3 positifs et de somme 1 tel que F puisse s'écrire sous la forme*

$$F = \alpha_1 \; F_1 + \alpha_2 \; F_2 + \alpha_3 \; F_3.$$

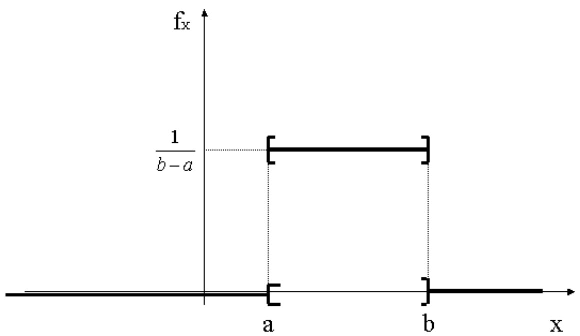
Loi uniforme

On dit qu'une v.a.r. X suit une **loi uniforme** sur un intervalle $[a, b]$, et on note $X \sim \mathcal{U}([a, b])$, si :

$$X(\Omega) = [a, b]$$

$$f(x) = \frac{1}{b-a} \, \mathbb{1}_{[a,b]}(x) \, .$$

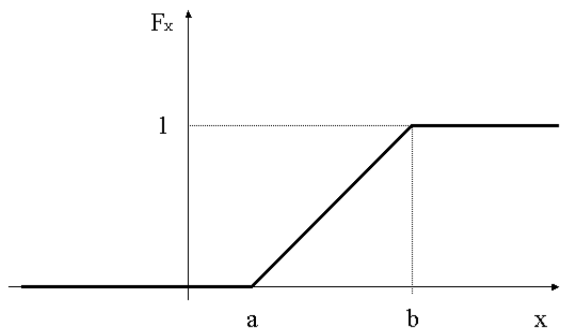
La densité de probabilité de la v.a.r. X a l'allure suivante :



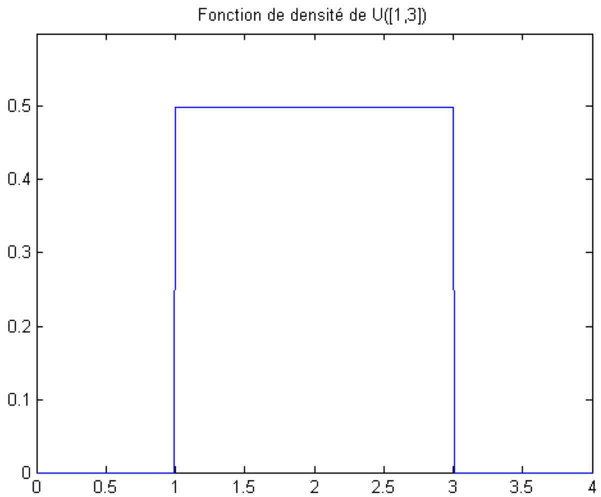
La fonction de répartition de la v.a.r. X vaut :

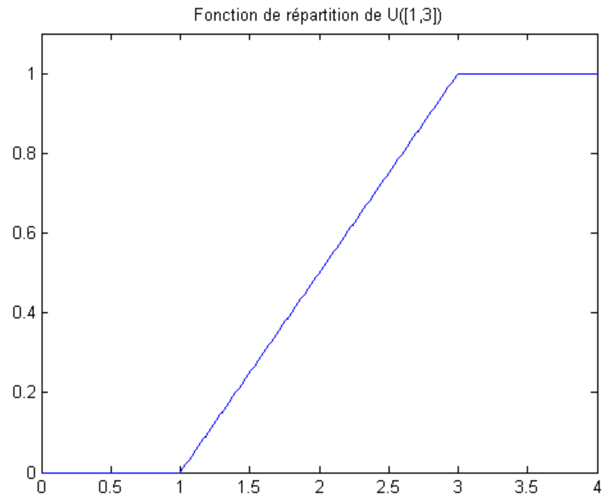
$$F(x) = \begin{cases} 0 & \text{si} \quad x \leq a \\ \dfrac{x-a}{b-a} & \text{si} \quad a \leq x \leq b \\ 1 & \text{si} \quad x \geq b \end{cases}$$

et a l'allure suivante :



Exemple


Fonction de densité de U([1,3])
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5
4


Fonction de répartition de U([1,3])
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
2.5
3
3.5
4

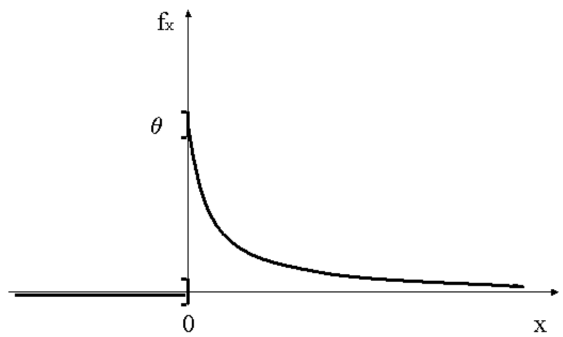
Loi exponentielle

On dit qu'une v.a.r. X suit une **loi exponentielle** de paramètre $\theta \in \mathbb{R}^{+*}$, et on note $X \sim \mathcal{E}(\theta)$, si :

$$X(\Omega) = \mathbb{R}^{+*}$$

$$f(x) = \theta e^{-\theta x} \; \mathbb{1}_{\mathbb{R}^{+*}}(x).$$

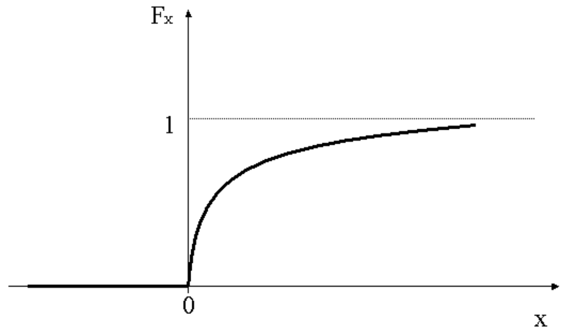
La densité de probabilité de la v.a.r. X a l'allure suivante :



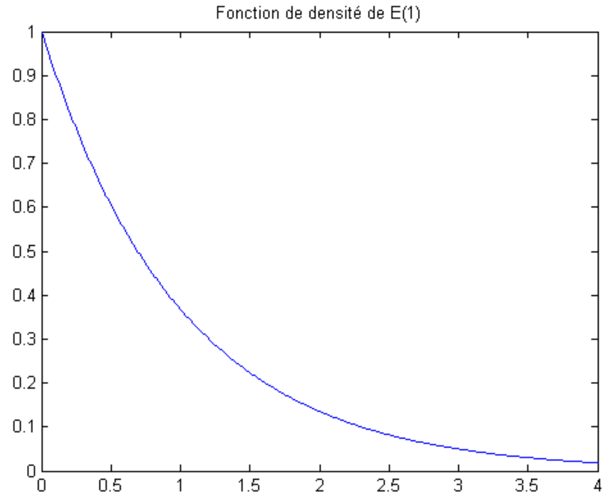
La fonction de répartition de la v.a.r. X vaut :

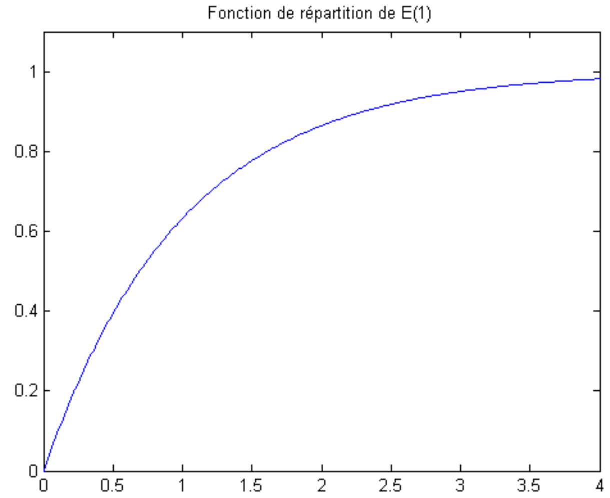
$$F(x) = \begin{cases} 0 & \text{si} \quad x \leq 0 \\ 1 - e^{-\theta x} & \text{si} \quad x \geq 0. \end{cases}$$

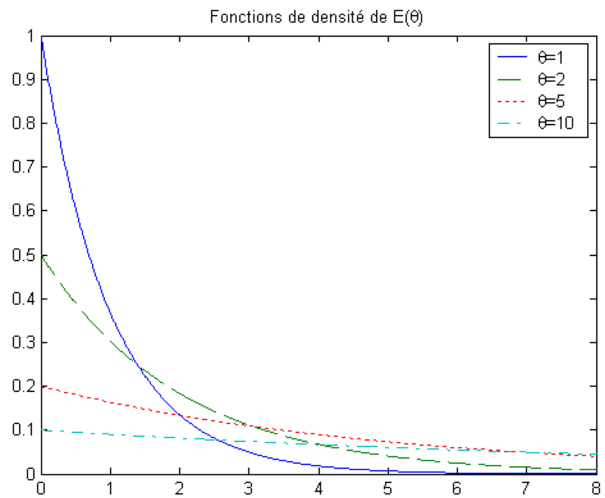
Elle a l'allure suivante :



Exemples


Fonction de densité de E(1)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.5
1.5
2
2.5
3
3.5
4


Fonction de répartition de E(1)
1
0.8
0.6
0.4
0.2
0
0.5
1.5
2
2.5
3
3.5
4


Fonctions de densité de E(θ)
θ=1
θ=2
θ=5
θ=10
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5
6
7
8

<u>Utilisation</u>

Cette loi est la plus utilisée dans le domaine de la fiabilité. Un système modélisé par la loi exponentielle est dit sans mémoire ou encore markovien.

Considérons par exemple un lot de machines identiques et indépendantes, et modélisons le temps d'attente avant l'arrivée d'une panne par une loi exponentielle de paramètre θ. On met en marche une machine et dès que survient une panne, on la remplace immédiatement par une autre machine, et ainsi de suite sur un intervalle de temps $[0, t]$. Le nombre de pannes observées durant cette période suit alors une loi de Poisson de paramètre θt.

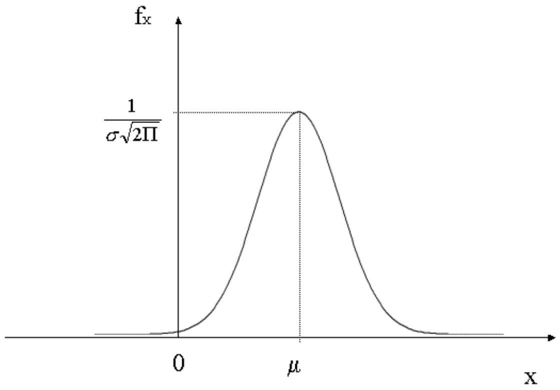
Loi normale

On dit qu'une v.a.r. X suit une **loi normale** (ou loi de Laplace-Gauss) de paramètres $\mu \in \mathbb{R}$ et σ^2 avec $\sigma \in \mathbb{R}^{+*}$, et on note $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, si :

$$X(\Omega) = \mathbb{R}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right).$$

La densité de probabilité de la v.a.r. X a l'allure suivante :



Remarques

1. On trouvera parfois la notation suivante : $X \sim \mathcal{N}(\mu, \sigma)$. Le type de notation choisi devra être clairement indiqué.
2. $\mathcal{N}(0,1)$ est appelée **loi normale centrée réduite**.
3. La fonction de répartition d'une loi normale $\mathcal{N}\left(\mu, \sigma^2\right)$ n'a pas d'expression explicite mais peut être exprimée en fonction de celle de la loi normale centrée réduite :

   $$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$

   Cette fonction est tabulée sous forme d'abaque.
   On notera φ la densité de probabilité de la loi normale centrée réduite.
4. $$X \sim \mathcal{N}\left(\mu, \sigma^2\right) \Rightarrow \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$$
   $$X \sim \mathcal{N}(0,1) \Rightarrow \sigma X + \mu \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

5. Si :

   (a) $X_1 \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right)$

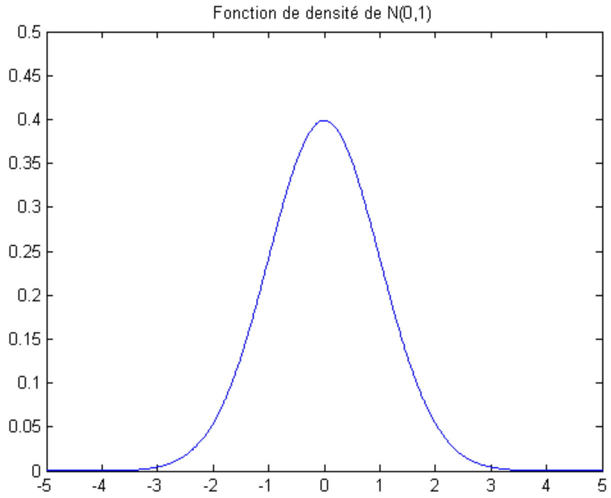
   (b) $X_2 \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right)$

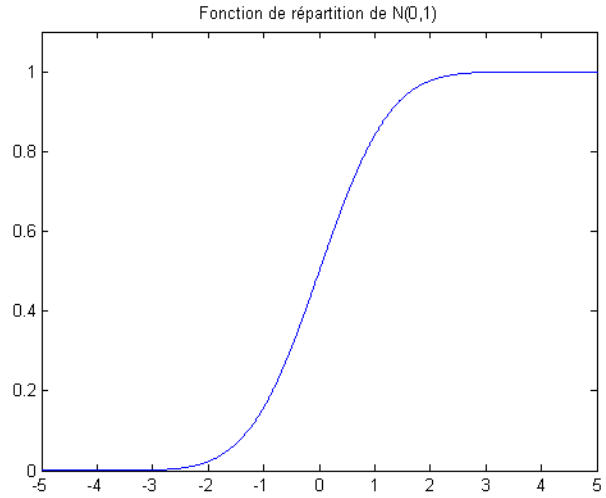
   (c) $X_1 \perp\!\!\!\perp X_2$

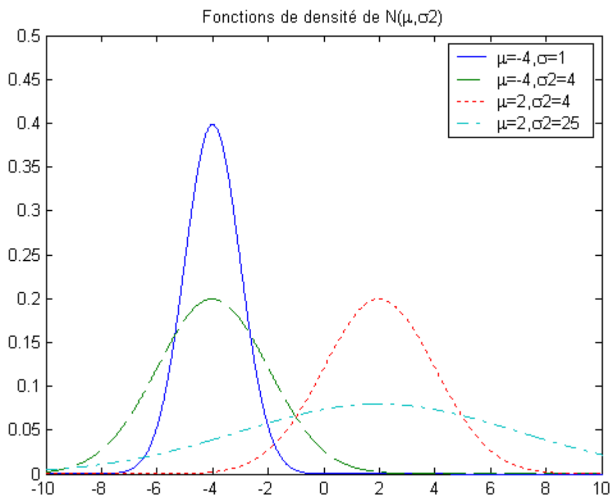
   alors on a :

$$X_1 + X_2 \sim \mathcal{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right).$$

Exemples


Fonction de densité de N(0,1)


Fonction de répartition de N(0,1)


Fonctions de densité de N(μ,σ2)
μ=-4,σ=1
μ=-4,σ2=4
μ=2,σ2=4
μ=2,σ2=25

Utilisation

Cette loi est la plus utilisée dans le domaine statistique. Elle sert notamment à modéliser les erreurs de mesure.

Loi log-normale

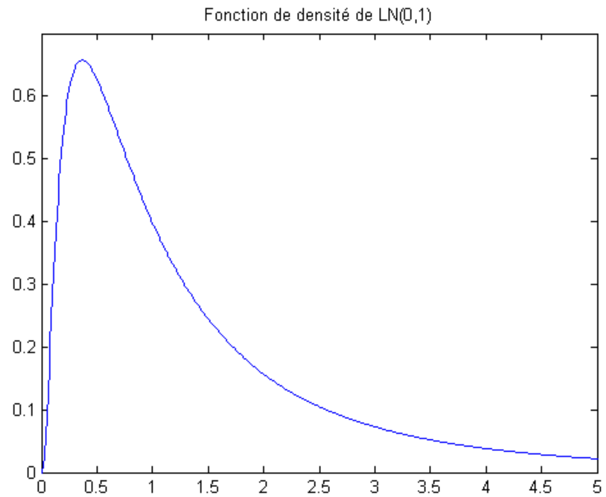
On dit qu'une v.a.r. X suit une **loi log-normale** de paramètres $\mu \in \mathbb{R}$ et σ^2 avec $\sigma \in \mathbb{R}^{+*}$, et on note $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$, si :

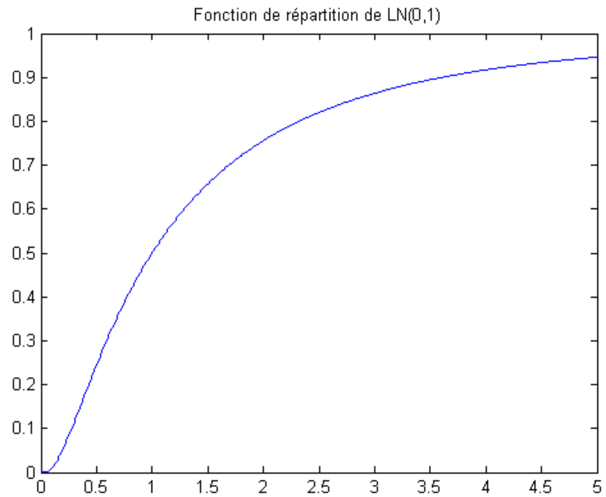
$$X(\Omega) = \mathbb{R}^{+*}$$

$$\log(X) \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

où log désigne le logarithme népérien (également noté ln).

Exemple





Utilisation

Cette loi est utilisée par exemple pour modéliser la taille des grains de sédiment provenant d'une couche géologique.

Loi Gamma

On dit qu'une v.a.r. X suit une **loi Gamma** de paramètres $p \in \mathbb{R}^{+*}$ et $\theta \in \mathbb{R}^{+*}$, et on note $X \sim \gamma(p, \theta)$, si :

$$X(\Omega) = \mathbb{R}^{+*}$$

$$f(x) = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1} \, \mathbb{1}_{\mathbb{R}^{+*}}(x).$$

La fonction gamma (Γ) est telle que :

$$\forall \alpha \in \mathbb{R} : \Gamma(\alpha) = \int_0^{+\infty} e^{-u} u^{\alpha-1} du.$$

On a :

$$\forall \alpha \in \mathbb{R} : \Gamma(\alpha+1) = \alpha\Gamma(\alpha)$$
$$\forall n \in \mathbb{N} : \Gamma(n+1) = n!$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Cas particuliers

1. Si $\theta = 1$ alors on note la loi $\gamma(p)$.
2. Si $p = 1$ alors il s'agit d'une loi exponentielle $\mathcal{E}(\theta)$.
3. Si $p = \dfrac{n}{2}$ et $\theta = \dfrac{1}{2}$ alors il s'agit d'une **loi du Khi-deux** de paramètre n, notée χ_n^2 :
$$\chi_n^2 = \gamma\left(\frac{n}{2}, \frac{1}{2}\right), \quad f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \, \mathbb{1}_{\mathbb{R}^{+*}}(x)$$

Remarques

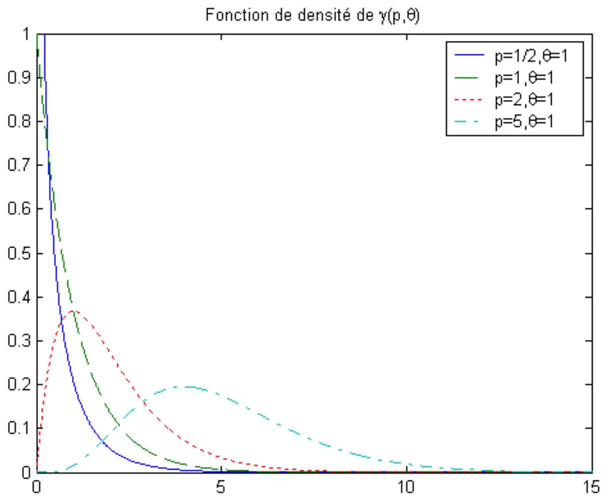
1. Si :
$$X_1 \sim \gamma(p_1, \theta), \quad X_2 \sim \gamma(p_2, \theta), \quad X_1 \perp\!\!\!\perp X_2,$$
alors on a :
$$X_1 + X_2 \sim \gamma(p_1 + p_2, \theta).$$

2. Si $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ alors $\left(\dfrac{X-\mu}{\sigma}\right)^2 \sim \chi_1^2$.

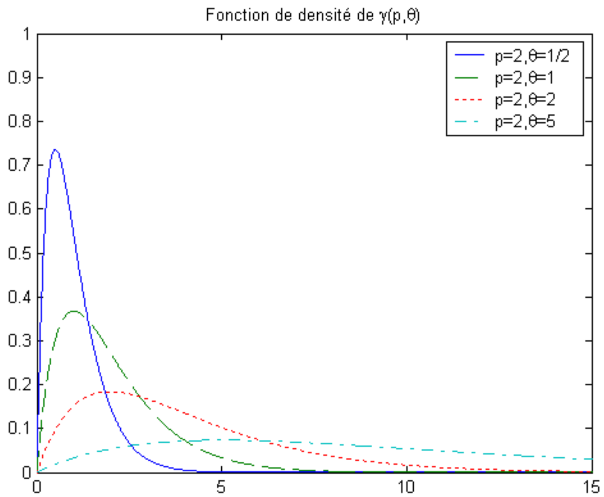
3. Si :
   (a) $\forall i \in \{1, ..., n\}$: $X_i \sim \mathcal{N}(0, 1)$
   (b) $(X_1, ..., X_n)$ indépendantes

   alors on a :
$$X_1^2 + ... + X_n^2 \sim \chi_n^2.$$

Exemples



Fonction de densité de γ(p,θ)
p=1/2,θ=1
p=1,θ=1
p=2,θ=1
p=5,θ=1



Fonction de densité de γ(p,θ)
p=2,θ=1/2
p=2,θ=1
p=2,θ=2
p=2,θ=5

Loi Beta de seconde espèce

On dit qu'une v.a.r. X suit une **loi Beta de seconde espèce** de paramètres $p \in \mathbb{R}^{+*}$ et $q \in \mathbb{R}^{+*}$, et on note $X \sim \beta_{II}(p, q)$, si :

$$X(\Omega) = \mathbb{R}^{+*}$$

$$f(x) = \frac{1}{\beta(p, q)} \frac{x^{p-1}}{(1 + x)^{p+q}} \mathbb{1}_{\mathbb{R}^{+*}}(x) .$$

La fonction beta (β) est telle que :

$$\beta(p, q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p + q)}.$$

Remarque

Si :

1. $X \sim \gamma(p)$
2. $Y \sim \gamma(q)$
3. $X \perp\!\!\!\perp Y$

alors :

$$\frac{X}{Y} \sim \beta_{II}(p, q).$$

Loi Beta de première espèce

On dit qu'une v.a.r. X suit une **loi Beta de première espèce** de paramètres $p \in \mathbb{R}^{+*}$ et $q \in \mathbb{R}^{+*}$, et on note $X \sim \beta_I(p,q)$, si :

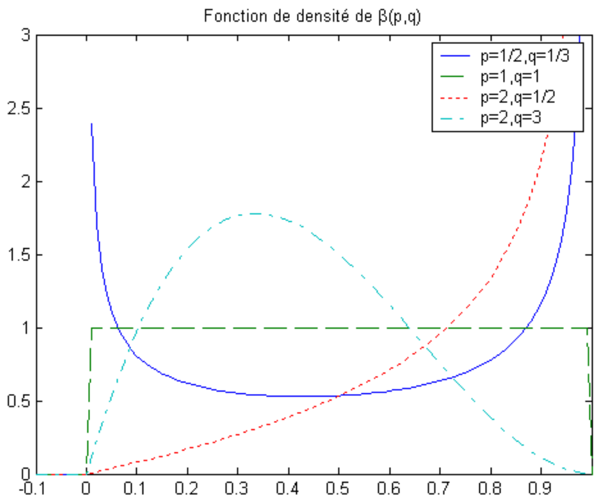
$$X(\Omega) = [0,1]$$

$$f(x) = \frac{1}{\beta(p,q)} x^{p-1} (1-x)^{q-1} \; \mathbb{1}_{[0,1]}(x).$$

Remarque

Si $X \sim \beta_{II}(p,q)$ alors $\dfrac{X}{1+X} \sim \beta_I(p,q)$.

Exemples



Fonction de densité de β(p,q)
p=1/2,q=1/3
p=1,q=1
p=2,q=1/2
p=2,q=3
3
2.5
2
1.5
1
0.5
0
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

Loi de Student

On dit qu'une v.a.r. X suit une **loi de Student** de paramètre $n \in \mathbb{N}^*$, et on note $X \sim \mathcal{T}(n)$, si :

$$X(\Omega) = \mathbb{R}$$

$$f(x) = \frac{1}{\sqrt{n}\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

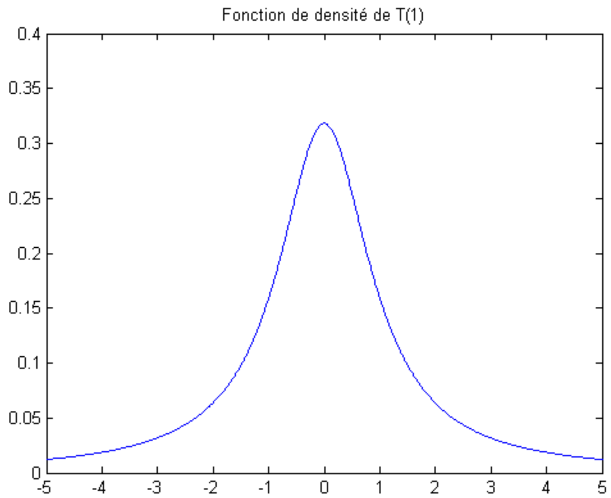
Remarque
Si :

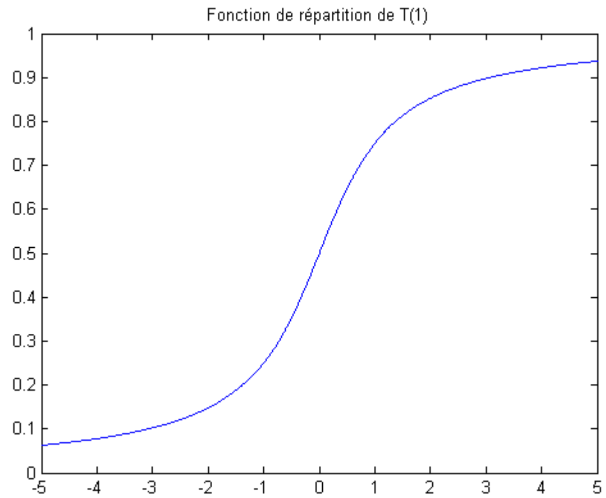
$$U \sim \mathcal{N}(0,1), \quad X \sim \chi_n^2, \quad U \perp\!\!\!\perp X,$$

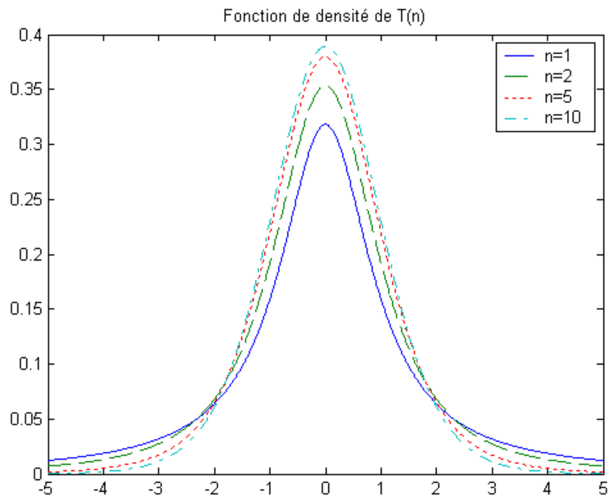
alors on a :

$$\frac{U}{\sqrt{\frac{X}{n}}} \sim \mathcal{T}(n).$$

Exemples


Fonction de densité de T(1)


Fonction de répartition de T(1)


Fonction de densité de T(n)
n=1
n=2
n=5
n=10

Loi de Fisher-Snedecor

On dit qu'une v.a.r. X suit une **loi de Fischer-Snedecor** de paramètres $m \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$, et on note $X \sim \mathcal{F}(m,n)$, si :

$$X(\Omega) = \mathbb{R}^+$$

$$f(x) = \frac{1}{\beta\left(\frac{m}{2}, \frac{n}{2}\right)} \, m^{\frac{m}{2}} n^{\frac{n}{2}} \, \frac{x^{\frac{m}{2}-1}}{(mx+n)^{\frac{m+n}{2}}} \, \mathbb{1}_{\mathbb{R}^+}(x).$$

Remarque

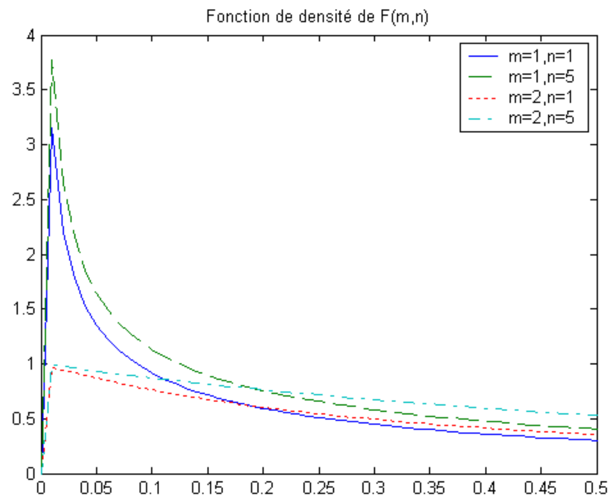
Si :

$$X \sim \chi^2_m, \quad Y \sim \chi^2_n, \quad X \perp\!\!\!\perp Y,$$

alors on a :

$$\frac{\frac{X}{m}}{\frac{Y}{n}} \sim \mathcal{F}(m,n).$$

Exemples



Remarque Les lois du Khi-deux, de Student et de Fisher-Snedecor sont très utilisées en Statistique, notamment dans le cadre des modèles linéaires (régression, analyse de variance).

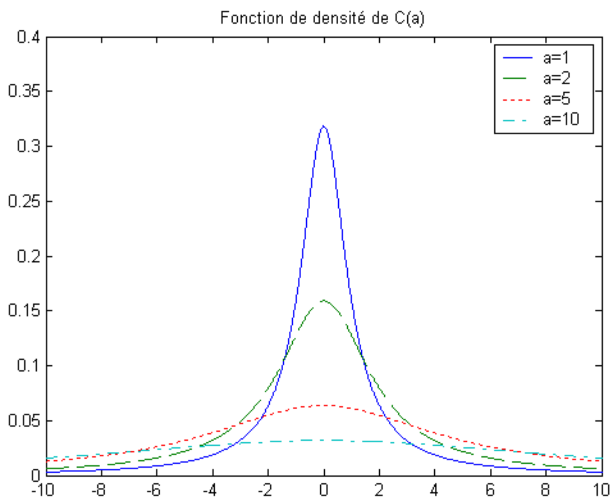
Loi de Cauchy

On dit qu'une v.a.r. X suit une **loi de Cauchy** de paramètre $a \in \mathbb{R}^{+*}$, et on note $X \sim \mathcal{C}(a)$, si :

$$X(\Omega) = \mathbb{R}$$

$$f(x) = \frac{a}{\pi} \frac{1}{x^2 + a^2}.$$

Exemples



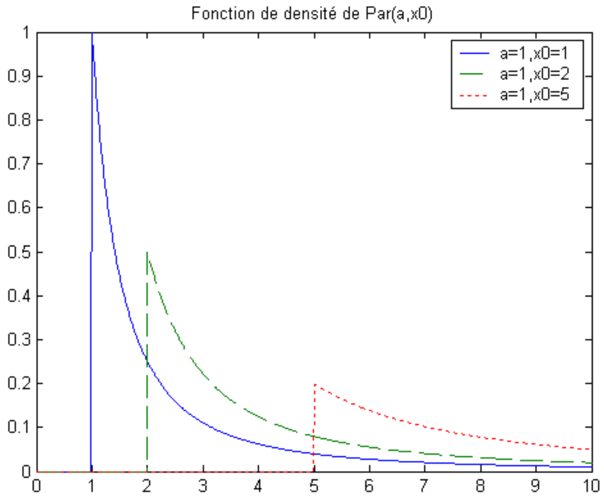
Loi de Pareto

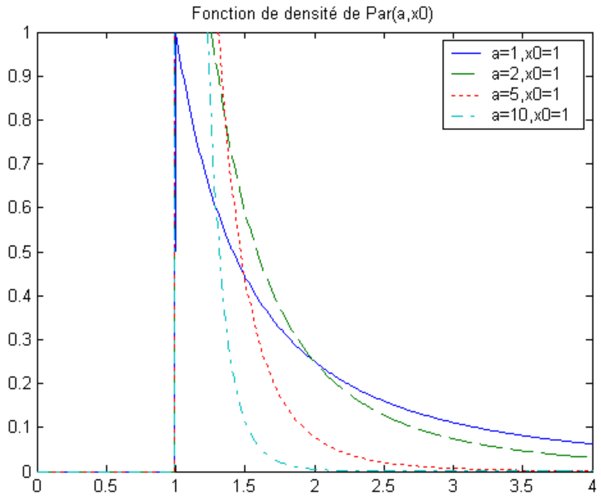
On dit qu'une v.a.r. X suit une **loi de Pareto** de paramètres $\alpha \in \mathbb{R}^{+*}$ et $x_0 \in \mathbb{R}^{+*}$, et on note $X \sim \mathcal{P}ar\left(\alpha, x_0\right)$, si :

$$X\left(\Omega\right) = [x_0, +\infty)$$

$$f\left(x\right) = \frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1} \cdot \mathbb{1}_{[x_0,+\infty)}\left(x\right).$$

Exemples


Fonction de densité de Par(a,x0)
a=1,x0=1
a=1,x0=2
a=1,x0=5


Fonction de densité de Par(a,x0)
a=1,x0=1
a=2,x0=1
a=5,x0=1
a=10,x0=1

2.1.3 Changement de variable

Théorème 2.1.15 *Soit U et V deux ouverts de $\mathbb{R}$ et X une v.a.r. absolument continue à valeurs dans U et de densité f_X. Soit φ une bijection de U vers $V = \operatorname{Im} \varphi$, continûment différentiable ainsi que son inverse (φ est dite $\mathcal{C}^1-$difféomorphisme). Alors la v.a.r. $Y = \varphi(X)$ est absolument continue, à valeurs dans V et de densité :*

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|\left(\varphi^{-1}\right)'(y)\right| \cdot \mathbb{1}_V(y).$$

2.2 Vecteurs aléatoires

2.2.1 Fonction de répartition

Définition 2.2.1 *On appelle* **fonction de répartition conjointe** *du vecteur aléatoire* $X = (X_1, \ldots, X_n)$ *la fonction* $F_X : \mathbb{R}^n \to [0, 1]$ *définie par :*

$$F_X(x) = P_X((-\infty, x_1] \times \ldots \times (-\infty, x_n]) = P\left(\bigcap_{i=1}^{n} \{X_i \leq x_i\}\right)$$

pour tout $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$*.*

Proposition 2.2.2 *On a :*

$$\lim_{\forall i \in \{1,\ldots,n\}: x_i \to -\infty} F_X(x) = 0, \qquad \lim_{\forall i \in \{1,\ldots,n\}: x_i \to +\infty} F_X(x) = 1.$$

Définition 2.2.3 *Tout sous vecteur du vecteur aléatoire X de dimension strictement inférieure à n, et différente de 1, est appelé* **vecteur aléatoire marginal** *(ou marge). Tout sous vecteur du vecteur aléatoire X de dimension 1 est appelé* **variable aléatoire marginale**.

Proposition 2.2.4 *La fonction de répartition conjointe d'un vecteur aléatoire $X = (X_1, \ldots, X_n)$ permet de déterminer les fonctions de répartition de toutes les marges.*

2.2.2 Densité de probabilité

Définition 2.2.5 *On dit que le vecteur aléatoire X (ou sa loi) est* **absolument continu(e)** *s'il existe une fonction mesurable*

$$f_X : (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) \to \left(\mathbb{R}^+, \mathcal{B}_{\mathbb{R}^+}\right)$$

telle que :

$$F_X(x) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(u_1, \dots, u_n)\ du_1 \dots du_n$$

pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. La fonction f_X est appelée **densité de probabilité conjointe** *du vecteur aléatoire X.*

Proposition 2.2.6 *Toute densité de probabilite conjointe f_X sur $\mathbb{R}^n$ vérifie :*

(a) f_X *est positive,*

(b) f_X *est intégrable et*

$$\int_{\mathbb{R}^n} f_X(x_1, \ldots, x_n)\; dx_1 \ldots dx_n = 1.$$

Réciproquement, toute fonction mesurable $f_X : (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *vérifiant les propriétés précédentes est une densité de probabilité.*

Remarque Si la densité de probabilité f_X est continue sur un ouvert $I \subset \mathbb{R}^n$, alors la fonction de répartition F_X est différentiable sur I et on a :

$$\forall a = (a_1, ..., a_n) \in I : f_X (a_1, ..., a_n) = \frac{\partial^n F_X}{\partial x_1 ... \partial x_n} (a_1, ..., a_n) .$$

Proposition 2.2.7 *Si X est un vecteur aléatoire absolument continu alors tout vecteur aléatoire marginal est également absolument continu et sa densité de probabilité est obtenue en intégrant la densité de probabilité conjointe de X par rapport aux coordonnées restantes.*

2.2.3 Loi conditionnelle et densité conditionnelle

Variables aléatoires discrètes

Définition 2.2.8 *Soit* $Z = (X, Y) \in \mathbb{R}^n \times \mathbb{R}^m$ *un vecteur aléatoire discret, de* **support** D *fini ou dénombrable :* $P_Z(\{z\}) > 0$ *pour tout* $z \in D$ *et* $P_Z\left(\overline{D}\right) = 0$. *Notons* I *et* J *les supports des lois* P_X *et* P_Y*, respectivement. Pour tout* $x \in I$*, la probabilité* $P_{Y/X=x}$ *définie sur* J *par*

$$P_{Y/X=x}(\{y\}) = P(Y = y/X = x) = \frac{P(Y = y, X = x)}{P(X = x)}, \quad \forall y \in J,$$

est appelée loi de probabilité de Y *conditionnelle à* $X = x$.

Variables aléatoires continues

Définition 2.2.9 *Soit* $Z = (X, Y) \in \mathbb{R}^n \times \mathbb{R}^m$ *un vecteur aléatoire de loi absolument continue, de densité* f_Z*. Soit* f_X *la densité de* X *et soit* $x \in A = \{x \in \mathbb{R}^n : f(x) > 0\}$ *(on appelle* **support** *de la loi* P_X *la fermeture de* A*, i.e. le plus petit fermé contenant* A*). La fonction*

$$f_{Y/X=x} \; : \; \mathbb{R}^m \longrightarrow \mathbb{R}^+$$

$$y \longmapsto \frac{f_Z(x, y)}{f_X(x)}$$

est une densité de probabilité sur $\mathbb{R}^m$ *et est appelée* **densité conditionnelle** *de* Y *sachant que* $X = x$*. On note* $P_{Y/X=x}$ *la loi de probabilité associée, appelée loi de* Y *sachant que* $X = x$*.*

2.2.4 Changement de variable

Théorème 2.2.10 *Soit X un vecteur aléatoire à valeurs dans un ouvert $U \subset \mathbb{R}^n$, de loi absolument continue, de densité f_X. Soit φ un $\mathcal{C}^1-$difféomorphisme de U vers $V = \operatorname{Im}\varphi$. Alors le vecteur aléatoire $Y = \varphi(X)$ est absolument continu, de densité :*

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|J_{\varphi^{-1}}(y)\right| \cdot \mathbb{1}_V(y), \quad \forall y \in \mathbb{R}^n,$$

où $J_{\varphi^{-1}}$ est le jacobien de φ^{-1}.

2.2.5 Indépendance

Théorème 2.2.12 *Soit $X = (X_1, ..., X_n)$ un vecteur aléatoire. Il y a équivalence entre les assertions suivantes :*

(a) $\coprod_{i=1}^{n} X_i$;

(b) $F_X = \prod_{i=1}^{n} F_{X_i}$;

(c) $f_X = \prod_{i=1}^{n} f_{X_i}$, *si X est un vecteur aléatoire absolument continu.*

Enfin, si $n = 2$, celles-ci sont encore équivalentes à chacune des deux suivantes :

(d) $f_{X_2/X_1=x_1} = f_{X_2}$;

(e) $f_{X_1/X_2=x_2} = f_{X_1}$.

Théorème 2.2.13 *Soient X et Y deux v.a.r. indépendantes et absolument continues de densités de probabilité respectives f_X et f_Y. La v.a.r. $X+Y$ est absolument continue et admet comme fonction de densité :*

$$f_{X+Y}(u) = \int f_X(u-v)\, f_Y(v)\, dv, \quad \forall u \in \mathbb{R}.$$

Il s'agit du **produit de convolution** *de f_X et f_Y :*

$$f_{X+Y} = f_X * f_Y.$$

2.3 Extension de la notion de densité

2.3.1 Intégrale par rapport à une mesure

Définition 2.3.1 *On appelle* **fonction étagée** *une fonction mesurable* f *définie sur un espace probabilisable* $(\Omega, \mathcal{A})$ *à valeurs dans* $\left(\overline{\mathbb{R}}, \mathcal{B}_{\overline{\mathbb{R}}}\right)$*, ne prenant qu'un nombre fini de valeurs* $(y_i)_{i \in \{1,\dots,n\}}$.

Théorème 2.3.2 *Toute fonction réelle positive est mesurable si elle est limite croissante d'une suite de fonctions étagées positives.*

On note $\mathcal{E}^+ = \mathcal{E}^+(\Omega, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables étagées positives définies sur l'espace mesuré $(\Omega, \mathcal{A}, \mu)$.

Définition 2.3.3 *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et soit $f \in \mathcal{E}^+$,*

$$f = \sum_{i=1}^{n} y_i \ \mathbb{1}_{A_i}.$$

On appelle **intégrale** *de f par rapport à μ le nombre (éventuellement infini) $\int f d\mu$ défini par :*

$$\int f d\mu = \sum_{i=1}^{n} y_i \ \mu(A_i).$$

Proposition 2.3.4 *Soit f, $g \in \mathcal{E}^+(\Omega, \mathcal{A}, \mu)$ et λ, $\gamma \in \mathbb{R}^+$. Alors :*

(a) $\lambda f + \gamma g \in \mathcal{E}^+(\Omega, \mathcal{A}, \mu)$ *et*

$$\int (\lambda f + \gamma g)\ d\mu = \lambda \int f\ d\mu + \gamma \int g\ d\mu;$$

(b) si $f \leq g$,

$$\int f\ d\mu \leq \int g\ d\mu.$$

Définition 2.3.5 *Soit f une fonction réelle, mesurable et positive. On appelle* **intégrale de f par rapport à** μ *le nombre* $\int f d\mu$ *défini par*

$$\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu,$$

où $\{f_n\}_n$ *est une suite de fonctions étagées positives croissantes vers* f.

Définition 2.3.6 *Une fonction réelle mesurable f est dite* $\mu-$**intégrable** *si*

$$\int f^+ \, d\mu < \infty \qquad et \qquad \int f^- \, d\mu < \infty,$$

où $f^+ = \max(f, 0)$ *et* $f^- = -\inf(f, 0)$. *Le réel*

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

est alors appelé **intégrale de f par rapport à** μ.

Proposition 2.3.7 *f est $\mu-$intégrable si et seulement si $\int |f| \ d\mu < \infty$.*

Exemples

1. Intégrale par rapport à une mesure de Dirac :
$$\int f \ d\delta_a = f(a).$$

2. Intégrale par rapport à une mesure discrète : si $\mu = \sum_{n \in I \subset \mathbb{N}} p_n \ \delta_{a_n}$,
$$\int f \ d\mu = \sum_{n \in I \subset \mathbb{N}} p_n \ f(a_n).$$

3. Intégrale par rapport à la mesure de Lebesgue sur $\mathbb{R}$:
$$\int_{[a,b]} f \ d\lambda = \int_a^b f(x) \ dx.$$

Proposition 2.3.8 *Propriétés de l'intégrale :*

(a) i. $\displaystyle\int (\lambda f + \gamma g) \ d\mu = \lambda \int f \ d\mu + \gamma \int g \ d\mu, \quad \forall \lambda, \mu \in \mathbb{R}$;

ii. $\displaystyle\int_{A \cup B} f \ d\mu = \int_A f \ d\mu + \int_B f \ d\mu, \quad \forall A, B \in \mathcal{A}, \ A \cap B = \emptyset$;

(b) $\displaystyle f \geq g \quad \Rightarrow \quad \int f \ d\mu \geq \int g \ d\mu$;

(c) $|f| \leq g$, g $\mu-$*intégrable* $\Rightarrow$ f $\mu-$*intégrable;*

(d) f $\mu-$*intégrable* $\displaystyle\Rightarrow \quad \left| \int f \ d\mu \right| \leq \int |f| \ d\mu.$

Définition 2.3.9 *Soit* $(\Omega, \mathcal{A}, \mu)$ *un espace mesuré et* $\mathcal{P}$ *une propriété définie sur* Ω*. On dira que* $\mathcal{P}$ *est vérifie* $\mu-$**presque partout** $(\mu -$ *p.p.*$)$ *s'il existe un ensemble* N *de* $\mathcal{A}$ *tel que* $\mu(N) = 0$ *et la propriété* $\mathcal{P}$ *est vérifie pour tout* ω *dans* $\overline{N}$*.*

Proposition 2.3.10 *(a)* $\int |f| \; d\mu < \infty \quad \Rightarrow \quad f < \infty \;\; \mu-$*p.p.*

(b) $f = 0 \;\; \mu-$*p.p.* $\iff \int |f| \; d\mu = 0 \quad \Rightarrow \quad \int f \; d\mu = 0$

(c) $\int |f| \; d\mu < \infty, \quad f = g \;\; \mu-$*p.p.*

$$\Rightarrow \quad \int |g| \; d\mu < \infty \quad et \quad \int f \; d\mu = \int g \; d\mu.$$

Théorème 2.3.11 *(Théorème de convergence monotone de Beppo-Levi)*
Soit $\{f_n\}_n$ *une suite croissante de fonctions mesurables positives. Alors*

$$\int \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Théorème 2.3.12 *(Théorème de convergence dominée de Lebesgue)*
Soit $\{f_n\}_n$ *une suite de fonctions réelles, mesurables telles que :*

(a) $f_n \underset{n\to+\infty}{\longrightarrow} f \quad \mu-$*p.p.*

(b) $\exists\, g : \Omega \to \overline{\mathbb{R}}^+, \quad \int g \, d\mu < \infty$ *telle que :*

$$|f_n| \leq g \ \mu - p.p., \quad \forall n \in \mathbb{N}.$$

Alors f_n *et* f *sont* $\mu-$*intégrables et*

$$\int f_n \, d\mu \underset{n\to+\infty}{\longrightarrow} \int f \, d\mu.$$

2.3.2 Continuité absolue d'une mesure par rapport à une autre. Densité

Définition 2.3.13 *Soit* $(\Omega, \mathcal{A})$ *un espace mesurable et* μ *et* ν *deux mesures positives sur cet espace.*

- *On dit que* ν **est absolument continue par rapport à** μ *si*

$$\forall A \in \mathcal{A}, \quad \mu(A) = 0 \quad \Longrightarrow \quad \nu(A) = 0.$$

On note $\nu \ll \mu$.

- *On dit que* ν **admet une densité par rapport à** μ *s'il existe une fonction mesurable positive* f *telle que*

$$\nu(A) = \int_A f \, d\mu, \quad \forall A \in \mathcal{A}.$$

Théorème 2.3.14 *Une mesure* ν *est absolument continue par rapport à une autre mesure* μ *si et seulement si* ν *admet une densité par rapport à* μ.

3 Moments de variables aléatoires

3.1 Variables aléatoires réelles intégrables et espérance mathématique

Définition 3.1.1 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* X *une v.a.r. définie sur* $(\Omega, \mathcal{A}, P)$ P*-intégrable :*

$$\int |X| \, dP < +\infty.$$

L'intégrale

$$\int X \; dP = \int_{\Omega} X(\omega) \; dP(\omega)$$

est appelée **espérance** *de la v.a.r.* X *est est notée* $E(X)$*.*

Théorème 3.1.2 (Théorème du transport)
Soit X *une v.a.r. de* $(\Omega, \mathcal{A}, P)$ *vers* $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *de loi* P_X *et* $h : \mathbb{R} \longrightarrow \mathbb{R}$ *une fonction* $\mathcal{B}_{\mathbb{R}}-$*mesurable, positive ou* P_X-*intégrable. On a alors :*

$$\begin{aligned}\int h(X)\ dP &= \int_\Omega (h \circ X)(\omega)\ dP(\omega) = \int_{\mathbb{R}} h(x)\ dP_X(x) \\ &= E[h(X)].\end{aligned}$$

Cas particuliers

1. Si X est une v.a.r. discrète à valeurs dans $D = \{x_n\}_{n \in I \subset \mathbb{N}}$ alors :

$$E(X) = \int X \, dP = \sum_{n \in I} x_n \, P(X = x_n).$$

2. Si X est une v.a.r. continue, absolument continue par rapport à la mesure de Lebesgue, de densité f, et intégrable alors :

$$E(X) = \int X \, dP = \int_{\mathbb{R}} x \, f(x) \, d\lambda(x) = \int_{-\infty}^{+\infty} x \, f(x) \, dx.$$

Exemples : lois discrètes

$$X \sim \delta_{\{a\}} : \quad E(X) = a$$

$$X \sim \mathcal{U}_{\{x_1,\dots,x_n\}} : \quad E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$X \sim \mathcal{B}(1,p) : \quad E(X) = p$$

$$X \sim \mathcal{B}(n,p) : \quad E(X) = np$$

$$X \sim \mathcal{G}(p) : \quad E(X) = \frac{1}{p}$$

$$X \sim \mathcal{BN}(n,p) : \quad E(X) = \frac{n}{p}$$

$$X \sim \mathcal{H}(n,N,M) : \quad E(X) = n\,\frac{M}{N}$$

$$X \sim \mathcal{P}(\lambda) : \quad E(X) = \lambda$$

Exemples : lois continues

$$X \sim \mathcal{U}_{[a,b]} : \quad E(X) = \frac{a+b}{2}$$

$$X \sim \mathcal{E}(\theta) : \quad E(X) = \frac{1}{\theta}$$

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right) : \quad E(X) = \mu$$

$$X \sim \mathcal{LN}\left(\mu, \sigma^2\right) : \quad E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$X \sim \gamma(p, \theta) : \quad E(X) = \frac{p}{\theta}$$

$$X \sim \beta_{II}(p, q) : \quad E(X) = \frac{p}{q-1} \quad \text{si } q \in (1, +\infty)$$

Si $q \in [0, 1]$, l'espérance n'existe pas.

$X \sim \beta_I(p,q)$: $\quad E(X) = \dfrac{p}{p+q}$

$X \sim \mathcal{T}(n)$: $\quad E(X) = 0$ si $n \in \mathbb{N} \setminus \{0,1\}$

$X \sim \mathcal{F}(m,n)$: $\quad E(X) = \dfrac{m}{m-2}$ si $m \in \mathbb{N} \setminus \{0,1,2\}$

$X \sim \mathcal{C}(a)$: $\quad E(X)$ n'existe pas, $E(|X|) = \infty$

$X \sim \mathcal{P}ar(\alpha, x_0)$: $\quad E(X) = \dfrac{\alpha}{\alpha - 1}\, x_0$ si $\alpha > 1$

Si $\alpha \leq 1$, l'espérance n'existe pas.

Proposition 3.1.3 (Inégalité de Jensen)
Si X une v.a.r. intégrable et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction mesurable convexe telle que $\varphi \circ X$ est intégrable, alors

$$\varphi\left(E\left(X\right)\right) \leq E\left(\varphi\left(X\right)\right).$$

Définition 3.1.4 *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. On note $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ l'ensemble des v.a.r. intégrables définies sur $(\Omega, \mathcal{A}, P)$.*

Définition 3.1.5 *On appelle $L^1(\Omega, \mathcal{A}, P)$ l'ensemble des classes d'équivalence sur $\mathcal{L}^1$.*

Théorème 3.1.6 *L'espace $L^1(\Omega, \mathcal{A}, P)$ est un espace vectoriel sur $\mathbb{R}$, normé par*

$$\|X\|_1 = \|X\|_{L^1} = E\left(|X|\right) = \int |X| \; dP.$$

3.2 Moments de variables aléatoires réelles

3.2.1 Espace L^p

Définition 3.2.1 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $p \geq 1$*. On appelle* $L^p(\Omega, \mathcal{A}, P)$ *l'ensemble des (classes d'équivalence de) v.a.r.* X *définies sur* $(\Omega, \mathcal{A}, P)$ *de puissance* p*-intégrable, c'est à dire telles que :*

$$E(|X|^p) < +\infty.$$

L'espace $L^p(\Omega, \mathcal{A}, P)$ est un espace vectoriel normé muni de la norme $\|X\|_p = \|X\|_{L^p} = [E(|X|^p)]^{\frac{1}{p}}$.

Définition 3.2.2 *Soit $X \in L^p(\Omega, \mathcal{A}, P)$. On définit :*

(a) moment d'ordre p : $E(X^p)$;

(b) moment absolu d'ordre p : $E(|X|^p)$;

(c) moment centré d'ordre p : $E[(X - E(X))^p]$.

Proposition 3.2.3 *Soient $p, q \in \mathbb{R}$ tels que $1 \leq p < q$. Alors*

$$\|X\|_{L^p} \leq \|X\|_{L^q}$$

et donc

$$L^q(\Omega, \mathcal{A}, P) \subset L^p(\Omega, \mathcal{A}, P).$$

Théorème 3.2.4 (Inégalité de Hölder)
Soit $p > 1$ *et* $q > 1$ *tels que* $\dfrac{1}{p} + \dfrac{1}{q} = 1$. *Soit* $X \in L^p(\Omega, \mathcal{A}, P)$ *et* $Y \in L^q(\Omega, \mathcal{A}, P)$. *Alors :*

$$\|XY\|_1 \leq \|X\|_p \cdot \|Y\|_q .$$

Théorème 3.2.5 (Inégalité de Minkovski)
Soit $p \geq 1$ *et* $X, Y \in L^p(\Omega, \mathcal{A}, P)$. *Alors :*

$$X + Y \in L^p(\Omega, \mathcal{A}, P) \qquad et \qquad \|X + Y\|_p \leq \|X\|_p + \|Y\|_p .$$

3.2.2 Espace L^2

Proposition 3.2.6 (Inégalité de Cauchy-Schwartz)

$$X, Y \in L^2 \quad \Longrightarrow \quad \|XY\|_1 \leq \|X\|_2 \cdot \|Y\|_2$$

Proposition 3.2.7 *L'application*

$$\begin{aligned} L^2 \times L^2 &\longrightarrow \mathbb{R} \\ (X, Y) &\longmapsto \langle X, Y \rangle = E\left(XY\right) = \int XY \, dP \end{aligned}$$

est un produit scalaire.

Théorème 3.2.8 L^2 *est un espace de Hilbert.*

Définition 3.2.9 *On appelle* **variance** *d'une v.a.r.* X *dans* L^2 *et on note* $V(X)$*, le moment centré d'ordre 2 de* X *:*

$$V(X) = E\left[(X - E(X))^2\right] = \langle X - E(X), X - E(X)\rangle .$$

On appelle **écart-type** *de* $X \in L^2$ *et on note* σ_X*, la racine carrée de sa variance :*

$$\sigma_X = \sqrt{V(X)}.$$

Proposition 3.2.10 *Pour* $X \in L^2$ *on a :*

(a) $V(X) = E\left(X^2\right) - [E(X)]^2$;

(b) $V(X) = 0 \quad \Leftrightarrow \quad X = a \in \mathbb{R} \;\; p.s.$;

(c) $V(aX + b) = a^2\, V(X)\,, \quad \forall a, b \in \mathbb{R}$;

(d) $V(X) = E\left[(X - E(X))^2\right] \;=\; \inf_{a \in \mathbb{R}} \left(E\left[(X - a)^2\right]\right)$.

Terminologie

Si $X \in L^2$, on appelle

1. $X - E(X)$: v.a.r. centrée associée à X;
2. $\dfrac{X - E(X)}{\sigma_X}$: v.a.r. centrée réduite associée à X.

Proposition 3.2.11

(a) **Inégalité de Markov** : *si* $X \in L^1$, *alors*

$$P(|X| \geq \varepsilon) \leq \frac{E(|X|)}{\varepsilon}, \quad \forall \varepsilon > 0.$$

(b) **Inégalité de Bienaymé-Tchebycheff** : *si* $X \in L^2$, *alors*

$$P(|X - E(X)| \geq \varepsilon) = \frac{V(X)}{\varepsilon^2}, \quad \forall \varepsilon > 0.$$

Proposition 3.2.12 *Soit $X, Y \in L^1$ deux v.a.r. indépendantes. Alors :*

(a) $XY \in L^1$,

(b) $E(XY) = E(X)\ E(Y)$.

Exemples : lois discrètes

$X \sim \delta_{\{a\}}$: $\quad V(X) = 0$

$X \sim \mathcal{U}_{\{x_1,\dots,x_n\}}$: $\quad V(X) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^2 - \left(\dfrac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

$X \sim \mathcal{B}(1,p)$: $\quad V(X) = p(1-p) = pq$

$X \sim \mathcal{B}(n,p)$: $\quad V(X) = np(1-p) = npq$

$X \sim \mathcal{G}(p)$: $\quad V(X) = \dfrac{1-p}{p^2}$

$X \sim \mathcal{BN}(n,p)$: $\quad V(X) = \dfrac{n(1-p)}{p^2}$

$X \sim \mathcal{H}(n,N,M)$: $\quad V(X) = \dfrac{N-n}{N-1}\, n\, \dfrac{M}{N} \left(1 - \dfrac{M}{N}\right)$

$X \sim \mathcal{P}(\lambda)$: $\quad V(X) = \lambda$

Exemples : lois continues

$$X \sim \mathcal{U}_{[a,b]} : \quad V(X) = \frac{(b-a)^2}{12}$$

$$X \sim \mathcal{E}(\theta) : \quad V(X) = \frac{1}{\theta^2}$$

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right) : \quad V(X) = \sigma^2$$

$$X \sim \mathcal{LN}\left(\mu, \sigma^2\right) : \quad V(X) = \left[\exp\left(\sigma^2\right) - 1\right] \exp\left(2\mu + \sigma^2\right)$$

$$X \sim \gamma(p, \theta) : \quad V(X) = \frac{p}{\theta^2}$$

$$X \sim \beta_{II}(p, q) : \quad V(X) = \frac{p(p+q-1)}{(q-1)^2(q-2)} \quad \text{si } q \in (2, +\infty)$$

Si $q \in [0, 2]$, la variance n'existe pas.

$X \sim \beta_I(p,q)$: $\quad V(X) = \dfrac{pq}{(p+q)^2(p+q+1)}$

Si $q \in [0,1]$, l'espérance n'existe pas.

$X \sim \mathcal{T}(n)$: $\quad V(X) = \dfrac{n}{n-2}$ si $n \in \mathbb{N} \backslash \{0,1,2\}$

Si $n \in \{1,2\}$, la variance n'existe pas (si $n = 0$ la loi n'est pas définie).

$X \sim \mathcal{F}(m,n)$: $\quad V(X) = \dfrac{2n^2}{m} \dfrac{m+n-2}{(n-2)^2(n-4)}$ si $n > 4$

Si $n \leq 4$, la variance n'existe pas.

$X \sim \mathcal{C}(a)$: $\quad V(X)$ n'existe pas (l'espérance n'existant pas)

$X \sim \mathcal{P}ar(\alpha, x_0)$: $\quad V(X) = \dfrac{\alpha^2 \; x_0^2}{(\alpha-1)^2 \; (a-2)}$ si $\alpha > 2$

Si $\alpha \leq 2$, la variance n'existe pas.

3.3 Vecteurs aléatoires

3.3.1 Espérance mathématique

Définition 3.3.1 *Soit* $X = (X_1, \ldots, X_n)^T$ *un vecteur aléatoire réel tel que* $X_i \in L^1$ *pour tout* $i \in \{1, \ldots, n\}$. *On appelle* **espérance** *du vecteur aléatoire* X *le vecteur* $E(X)$ *de composantes* $E(X_1), \ldots, E(X_n)$ *:*

$$E(X) = (E(X_1), \ldots, E(X_n))^T.$$

3.3.2 Covariance de deux v.a.r.

Définition 3.3.2 *Soit* X *et* Y *deux v.a.r. intégrables telles que* XY *soit intégrable. On appelle* **covariance** *de* X *et* Y *et on note* $Cov(X, Y)$ *le réel*

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Proposition 3.3.3 *Soit* $X,\ Y \in L^1$ *telles que* $XY \in L^1$.

(a) La covariance est un opérateur bilinéaire symétrique :

i. $Cov\left(Y, X\right) = Cov\left(X, Y\right)$,

ii. $Cov\left(\alpha X_1 + \beta X_2, Y\right) = \alpha\, Cov\left(X_1, Y\right) + \beta\, Cov\left(X_2, Y\right)$, $\forall \alpha, \beta \in \mathbb{R}$;

(b) $Cov\left(X, Y\right) = E\left(XY\right) - E\left(X\right)\ E\left(Y\right)$.

Proposition 3.3.4 *Si* $(X_i)_{i\in\{1,...,n\}} \subset L^2$*, alors :*

(a) $Cov\left(X_i, X_i\right) = V\left(X_i\right), \quad \forall i \in \{1, ..., n\}$;

(b) $V\left(X_i + X_j\right) = V\left(X_i\right) + V\left(X_j\right) + 2Cov\left(X_i, X_j\right), \quad \forall i \neq j$;

(c) $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V\left(X_i\right) + 2 \sum_{1\leq i<j\leq n} Cov\left(X_i, X_j\right)$.

Proposition 3.3.5 *Soit X, $Y \in L^1$ telles que $XY \in L^1$. Alors :*

$$X \perp\!\!\!\perp Y \qquad \Longrightarrow \qquad Cov\,(X,Y) = 0.$$

Corollaire 3.3.6 $\quad (X_i)_{i\in\{1,\dots,n\}} \subset L^2$

$$\perp\!\!\!\perp_{i=1}^{n} X_i \qquad \Longrightarrow \qquad V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V\,(X_i)$$

Définition 3.3.7 *Soit X et Y deux v.a.r. non constantes dans $L^2\,(\Omega, \mathcal{A}, P)$. On appelle* **coefficient de corrélation** *de X et Y et on note $\rho_{X,Y}$ le réel*

$$\rho_{X,Y} = \frac{Cov\,(X,Y)}{\sigma_X\ \sigma_Y} = \frac{Cov\,(X,Y)}{\sqrt{V\,(X)\ V\,(Y)}}.$$

Proposition 3.3.8 $\quad \rho_{X,Y} \in [-1, 1]\,, \quad \forall X, Y \in L^2$ *v.a.r. non constantes.*

3.3.3 Matrice de covariance

Définition 3.3.9 *Soit* $X = (X_1, \ldots, X_n)^T$ *un vecteur aléatoire réel tel que* $(X_i)_{i\in\{1,\ldots,n\}} \subset L^2$*. On appelle* **matrice de variance-covariance** *de* X *la matrice carrée (symétrique) de dimension* n*, notée* $V(X)$*, de terme général* $Cov\left(X_i, X_j\right)$ *:*

$$V(X) = \begin{bmatrix} V(X_1) & Cov(X_1, X_2) & ... & ... \\ Cov(X_2, X_1) & V(X_2) & ... & ... \\ ... & ... & ... & ... \\ ... & ... & ... & V(X_n) \end{bmatrix}$$

$$= \left(Cov\left(X_i, X_j\right)\right)_{1\leq i,j\leq n}$$

$$= E\left[(X - E(X))(X - E(X))^T\right].$$

Proposition 3.3.10 *Soit $X = (X_1, ..., X_n)^T$ un vecteur aléatoire réel de carré intégrable : $E\left(X^T X\right) < \infty$. Si A est une matrice réelle $p \times n$, alors :*

(a) $E(AX) = A\, E(X)$,

(b) $V(AX) = A\, V(X)\ A^T$.

3.3.4 Espérance conditionnelle

Définition 3.3.11 *Soit (X, Y) un vecteur de v.a.r. Si l'intégrale*

$$\int_{\mathbb{R}} y \, dP_{Y/X=x}(y)$$

existe, on l'appelle **espérance conditionnelle** *de Y sachant que $X = x$. On la note $E(Y/X = x)$.*

4 Caractérisation des lois : transformée de Laplace et fonction caractéristique

4.1 Transformée de Laplace

4.1.1 Variables aléatoires réelles

Définition 4.1.1 *Soit X une v.a.r. définie sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$*. Soit I un intervalle contenant* 0*, tel que, pour tout $s \in I$, la v.a.r. e^{sX} soit intégrable. On appelle* **transformée de Laplace** *de X la fonction $\mathcal{L}_X : I \to \mathbb{R}^+$ définie par :*

$$\mathcal{L}_X(s) = E\left(e^{sX}\right).$$

Proposition 4.1.2

(a) $\mathcal{L}_X(0)$ *existe toujours et* $\mathcal{L}_X(0) = 1$.

(b) Si X *est bornée, alors* $\mathcal{L}_X$ *est définie et continue sur* $\mathbb{R}$.

(c) Si X *est positive, alors* $\mathcal{L}_X$ *est continue et bornée sur* $(-\infty, 0]$.

Théorème 4.1.3 *La transformée de Laplace d'une v.a.r. caractérise la loi de cette v.a.r. Ainsi, si deux v.a.r. ont la même transformée de Laplace alors elles ont la même loi.*

Théorème 4.1.4 *Soient* X *et* Y *deux v.a.r. indépendantes, admettant respectivement une transformée de Laplace sur* I *et* J. *La somme de ces deux v.a.r. admet alors une transformée de Laplace sur* $I \cap J$ *d'expression :*

$$\mathcal{L}_{X+Y}(t) = \mathcal{L}_X(t)\,\mathcal{L}_Y(t), \quad \forall t \in I \cap J.$$

Théorème 4.1.5 *Soit X une v.a.r. définie sur $(\Omega, \mathcal{A}, P)$, admettant une transformée de Laplace sur $I = (-a, b)$ avec $a, b > 0$. On a alors :*

(a) $E\left(|X|^k\right) < +\infty, \quad \forall k \in \mathbb{N};$

(b) $\mathcal{L}_X(s) = \sum_{n=1}^{+\infty} \frac{s^n}{n!} E(X^n), \quad \forall s \in (-t, t)$, *où* $t = \min(a, b)$;

(c) $E\left(X^k\right) = \mathcal{L}_X^{(k)}(0), \quad \forall k \in \mathbb{N}.$

4.1.2 Vecteurs aléatoires

Définition 4.1.6 *Soit* $X = (X_1, \ldots, X_n)^T$ *un vecteur aléatoire défini sur* $(\Omega, \mathcal{A}, P)$ *et à valeur dans* $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$*. Soit* $I \subset \mathbb{R}^n$ *un ouvert contenant* $0 \in \mathbb{R}^n$*, tel que la v.a.r.* $e^{t^T X}$ *soit intégrable pour tout* $t \in I$*. On appelle* **transformée de Laplace** *de* X *la fonction* $\mathcal{L}_X : I \longrightarrow \mathbb{R}^+$ *définie par :*

$$\mathcal{L}_X(t) = E\left(e^{t^T X}\right) = E\left(e^{\langle t, X\rangle}\right) = E\left(\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right), \quad \forall t \in I.$$

Proposition 4.1.7 *Soit* $X = (X_1, ..., X_n)^T$ *un vecteur aléatoire de* $\mathbb{R}^n$*, admettant une transformée de Laplace sur un ouvert* $\mathcal{O} \subset \mathbb{R}^n$*. Les v.a.r.* $(X_i)_{i \in \{1, \ldots, n\}}$ *sont indépendantes si et seulement si :*

$$\mathcal{L}_X(t_1, \ldots, t_n) = \prod_{i=1}^{n} \mathcal{L}_{X_i}(t_i), \quad \forall t = (t_1, ..., t_n) \in \mathcal{O}.$$

4.2 Fonction caractéristique

4.2.1 Intégrale d'une variable aléatoire complexe

Définition 4.2.1 *On appelle variable aléatoire complexe toute application*

$$Z : (\Omega, \mathcal{A}, P) \to \mathbb{C}$$
$$\omega \mapsto Z(\omega) = X(\omega) + i\, Y(\omega)$$

où X et Y sont deux v.a.r. définies sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$.

Définition 4.2.2 *La v.a complexe Z est $P-$intégrable si et seulement si X et Y le sont. Dans ce cas on définit :*

$$E(Z) = \int Z\, dP = \int X\ dP + i \int Y\ dP = E(X) + i\ E(Y).$$

4.2.2 Fonction caractéristique

Définition 4.2.3 *Soit* $X = (X_1, \ldots, X_n)^T$ *un vecteur aléatoire de* $\mathbb{R}^n$*. On appelle* **fonction caractéristique** *de* X *la fonction*

$$\begin{aligned} \varphi_X : \mathbb{R}^n &\longrightarrow \mathbb{C} \\ t &\longmapsto \varphi_X(t) = E\left(e^{it^T X}\right) = \int e^{it^T x}\, dP_X(x) \end{aligned}$$

avec $x = (x_1, \ldots, x_n)^T$ *et* $t = (t_1, \ldots, t_n)^T$*.*

Théorème 4.2.4 *Soit* $X = (X_1, \ldots, X_n)^T$ *un vecteur aléatoire de* $\mathbb{R}^n$. *Soit* φ_X *la fonction caractéristique de* X *et soit* φ_{X_i} *la fonction caractéristique de* X_i, $1 \leq i \leq n$. *On a alors :*

(a) $|\varphi_X(t)| \leq \varphi_X(0) = 1, \quad \forall t \in \mathbb{R}^n$;

(b) $\varphi_{X_i}(t_i) = \varphi_X(0, ..., 0, t_i, 0, ..., 0), \quad \forall i \in \{1, \ldots, n\}, \quad \forall t_i \in \mathbb{R}$;

(c) $\overline{\varphi_X(t)} = \varphi_X(-t), \quad \forall t \in \mathbb{R}^n$;

(d) φ *est uniformément continue sur* $\mathbb{R}^n$;

(e) $A \in \mathbb{R}^{p \times n},\ b \in \mathbb{R}^p : \varphi_{AX+b}(t) = \exp\left(it^T b\right)\ \varphi_X\left(A^T t\right), \quad \forall t \in \mathbb{R}^n$;

(f) si $X \in L^p$, *avec* $p \in \mathbb{N}$ *alors* φ_X *est* p *fois dérivable et*

$$\varphi_X^{(k)}(0) = i^k\ E\left(X^k\right), \quad \forall k \leq p.$$

Théorème 4.2.5 *La fonction caractéristique caractérise la loi d'une variable aléatoire. Ainsi, si deux variables aléatoires ont la même fonction caractéristique alors elles suivent la même loi.*

Théorème 4.2.6 *Soit* $\varphi : \mathbb{R} \to \mathbb{C}$ *une fonction intégrable :*

$$\int |\varphi(t)| \; dt < +\infty.$$

Si la fonction f *définie par*

$$f(x) = \frac{1}{2\pi} \int \varphi(t) \; e^{-itx} \; dt$$

est également intégrable, alors φ *est la fonction caractéristique de la v.a.r.* X *ayant* f *pour densité de probabilité par rapport à la mesure de Lebesgue.*

Proposition 4.2.7 *Si X et Y sont deux v.a.r. indépendantes, alors*

$$\varphi_{X+Y}(t) = \varphi_X(t)\ \varphi_Y(t), \quad \forall t \in \mathbb{R}.$$

Proposition 4.2.8 *Les v.a.r. $(X_i)_{i\in\{1,\dots,n\}}$ sont indépendantes si et seulement si*

$$\varphi_X(t) = \prod_{i=1}^{n} \varphi_{X_i}(t_i), \quad \forall t = (t_1, \dots, t_n)^T \in \mathbb{R}^n,$$

où $X = (X_1, \dots, X_n)^T$.

5 Vecteurs gaussiens

5.1 Exemple fondamental

$X = (X_1, \ldots, X_n)^T : \quad X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right), \quad \forall i \in \{1, \ldots, n\}$

$\{X_1, \ldots, X_n\}$ famille de v.a.r. indépendantes

Alors :

- $E(X) = \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad V(X) = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix},$
- $f_X(x) = \dfrac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\mathbf{\Sigma})}} \exp\left[-\dfrac{1}{2}(x-\mu)^T \mathbf{\Sigma}^{-1}(x-\mu)\right], \quad \forall x \in \mathbb{R}^n,$
- $\varphi_X(t) = \exp\left(it^T\mu - \frac{1}{2}t^T\mathbf{\Sigma}t\right), \quad \forall t \in \mathbb{R}^n.$

5.2 Définition

Définition 5.2.1 *On dit qu'un vecteur aléatoire* $X = (X_1, \ldots, X_n)^T$ *de* $\mathbb{R}^n$ *est un* **vecteur gaussien** *si toute combinaison linéaire des composantes de* X *suit une loi normale :*

$$\lambda^T X = \sum_{i=1}^{n} \lambda_i \, X_i \sim \mathcal{N}\left(\mu_\lambda, \boldsymbol{\Sigma}_\lambda\right), \quad \forall \lambda = (\lambda_1, \ldots, \lambda_n)^T \in \mathbb{R}^n.$$

Si $E(X) = \mu$ *et* $V(X) = \boldsymbol{\Sigma}$*, on note* $X \sim \mathcal{N}(\mu, \boldsymbol{\Sigma})$*.*

Proposition 5.2.2

$$X \sim \mathcal{N}(\mu, \boldsymbol{\Sigma}) \qquad \Longrightarrow \qquad \lambda^T X \sim \mathcal{N}\left(\lambda^T \mu, \lambda^T \boldsymbol{\Sigma} \lambda\right), \quad \forall \lambda \in \mathbb{R}^n$$

Proposition 5.2.3 *Pour que le vecteur aléatoire X de $\mathbb{R}^n$ soit gaussien, il faut et il suffit qu'il existe un vecteur $\mu \in \mathbb{R}^n$ et une matrice symétrique et semi-définie positive $\mathbf{\Sigma}$ de dimension $n \times n$ tels que la fonction caractéristique s'écrive :*

$$\varphi_X(t) = \exp\left(it^T\mu - \frac{1}{2}t^T\mathbf{\Sigma}t\right), \quad \forall t \in \mathbb{R}^n.$$

Dans ce cas $X \sim \mathcal{N}(\mu, \mathbf{\Sigma})$.

Proposition 5.2.4 *Si $X \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ avec $\mathbf{\Sigma}$ inversible alors X est un vecteur gaussien non dégénéré et sa loi est absolument continue par rapport à la mesure de Lebesgue de $\mathbb{R}^n$; il admet comme densité de probabilité :*

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\mathbf{\Sigma})}} \exp\left[-\frac{1}{2}(x-\mu)^T\mathbf{\Sigma}^{-1}(x-\mu)\right], \quad \forall x \in \mathbb{R}^n.$$

5.3 Propriétés des vecteurs aléatoires gaussiens

5.3.1 Transformation linéaire d'un vecteur gaussien

Proposition 5.3.1 *Soit A une matrice $p \times n$. Alors*

$$X \sim \mathcal{N}_n(\mu, \mathbf{\Sigma}) \quad \Rightarrow \quad AX \sim \mathcal{N}_p\left(A\mu, A\mathbf{\Sigma}A^T\right).$$

5.3.2 Vecteur gaussien et indépendance

Proposition 5.3.2 *Soit* $X = (X_1, \ldots, X_n)^T \sim \mathcal{N}(\mu, \Sigma)$. *Il y a équivalence entre les assertions suivantes :*

(a) $\perp\!\!\!\perp_{i=1}^{n} X_i$,

(b) Σ *diagonale, i.e.*

$$Cov\left(X_i, X_j\right) = 0, \quad \forall i, j \in \{1, ..., n\}, \ i \neq j.$$

Corollaire 5.3.3 *Si le couple* (X, Y) *est un vecteur gaussien, alors*

$$X \perp\!\!\!\perp Y \quad \Leftrightarrow \quad Cov(X, Y) = 0.$$

6 Convergences

6.1 Convergence en loi

6.1.1 Définition

Définition 6.1.1 $X_n, X \in \mathbb{R}^p$

$$X_n \underset{n \to \infty}{\overset{\mathcal{L}}{\longrightarrow}} X \quad \Leftrightarrow \quad E\left[h\left(X_n\right)\right] \underset{n \to \infty}{\longrightarrow} E\left[h\left(X\right)\right], \forall h \textit{ fct. réelle, cont. et bornée}$$

Théorème 6.1.2 (Slutsky)

$$X_n \overset{\mathcal{L}}{\longrightarrow} X, \quad g : \mathbb{R}^p \to \mathbb{R}^q \text{ cont} \qquad \Rightarrow \qquad g\left(X_n\right) \overset{\mathcal{L}}{\longrightarrow} g\left(X\right)$$

6.1.2 Caractérisation de la convergence en loi

Proposition 6.1.3

$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad F_{X_n}(x) \underset{n\to\infty}{\longrightarrow} F_X(x)$ *en tout point de continuité de* F_X

Théorème 6.1.4 *Si* $(X_n)_{n\in\mathbb{N}}$ *et* X *sont des vecteurs aléatoires absolument continus et si* $\left(f_{X_n}\right)_{n\in\mathbb{N}}$ *et* f_X *désignent respectivement les fonctions de densité des vecteurs aléatoires* $(X_n)_{n\in\mathbb{N}}$ *et* X *alors on a :*

$$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad f_{X_n}(x) \underset{n\to\infty}{\longrightarrow} f_X(x) \quad \lambda_p - p.p.$$

Théorème 6.1.5 (Théorème de Paul Lévy)

(a) Soit $\left(\varphi_{X_n}\right)_{n\in\mathbb{N}}$*,* φ_X *les fonctions caractéristiques des vecteurs aléatoires* $(X_n)_{n\in\mathbb{N}}$ *et* X*. Alors :*

$$X_n \xrightarrow{\mathcal{L}} X \quad \Rightarrow \quad \varphi_{X_n}(t) \to \varphi_X(t)\,, \quad \forall t \in \mathbb{R}^p;$$

(b)

$$\varphi_{X_n}(t) \to \varphi(t)\,,\ \forall t \in \mathbb{R}^p,\ \varphi\ cont.\ en\ 0 \quad \Rightarrow \quad \varphi = \varphi_X\ \ et\ \ X_n \xrightarrow{\mathcal{L}} X.$$

Théorème 6.1.6 (Théorème de Cramer-Wold)

$$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad \lambda^T X_n \xrightarrow{\mathcal{L}} \lambda^T X, \quad \forall \lambda \in \mathbb{R}^p.$$

6.1.3 Approximation de lois

Proposition 6.1.7

$$X_n \sim \mathcal{B}(n, p_n), \quad np_n \underset{n\to\infty}{\longrightarrow} \lambda \quad \Rightarrow \quad X_n \xrightarrow{\mathcal{L}} X \sim \mathcal{P}(\lambda)$$

Théorème 6.1.8 (Théorème limite centrale)

$$(X_n)_n \ \in \ L^2, \ indépendants, \ E(X_n) = \mu, \ V(X_n) = \mathbf{\Sigma}, \ \forall n \in \mathbb{N}$$

$$\Rightarrow \quad \sqrt{n}\left(\overline{X}_n - \mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \mathbf{\Sigma})$$

Théorème 6.1.9 (Théorème limite centrale unidimensionnel)

$$(X_n)_n \ \in \ L^2, \ v.a.r. \ indépendantes, \ E(X_n) = \mu, \ V(X_n) = \sigma^2 > 0, \ \forall n \in \mathbb{N}$$

$$\Rightarrow \quad \sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

6.2 Convergence en probabilité

6.2.1 Définition

Définition 6.2.1 X_n, X *v.a.r.*

$$X_n \underset{n\to+\infty}{\overset{P}{\longrightarrow}} X \quad \Leftrightarrow \quad P\left(|X_n - X| > \varepsilon\right) \underset{n\to+\infty}{\longrightarrow} 0,\ \forall \varepsilon > 0$$

Proposition 6.2.2 *Soit* $(X_n)_n$ *une suite de v.a.r. dans* L^2*. Si on a*

$$\lim_{n\to\infty} E\left(X_n\right) = a, \qquad \lim_{n\to\infty} V\left(X_n\right) = 0,$$

alors

$$X_n \xrightarrow{P} a.$$

Proposition 6.2.3 (Loi faible des grands nombres)
Soit $(X_n)_n$ une suite de v.a.r. dans L^1, indépendantes et identiquement distribuées, de moyenne μ. Alors :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

Théorème 6.2.4 (Théorème de Slutsky)

$$X_n \xrightarrow{P} X, \quad g : \mathbb{R} \to \mathbb{R} \text{ cont} \qquad \Rightarrow \qquad g(X_n) \xrightarrow{P} g(X)$$

Définition 6.2.5 $X_n = \left(X_n^{(1)}, \ldots, X_n^{(p)}\right)^T, X = \left(X^{(1)}, \ldots, X^{(p)}\right)^T$ v.a.

$$X_n \xrightarrow[n\to\infty]{P} X \quad \Leftrightarrow \quad X_n^{(i)} \xrightarrow[n\to\infty]{P} X^{(i)}, \quad \forall i \in \{1, \ldots, n\}$$

Théorème 6.2.6 $X_n,\ X \in \mathbb{R}^p,\quad \|\cdot\|$ *norme sur* $\mathbb{R}^p$

$$X_n \xrightarrow[n\to\infty]{P} X \quad \Leftrightarrow \quad \|X_n - X\| \xrightarrow[n\to\infty]{P} 0$$

Proposition 6.2.7

$$X_n \xrightarrow{P} X, \quad Y_n \xrightarrow{P} Y, \quad g : \mathbb{R}^2 \to \mathbb{R} \text{ cont.}$$

$$\Rightarrow \quad g\left(X_n, Y_n\right) \xrightarrow{P} g\left(X, Y\right)$$

Corollaire 6.2.8

$$\left\{ \begin{array}{l} X_n \xrightarrow{P} X \\ \\ Y_n \xrightarrow{P} Y \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l} \alpha X_n + \beta Y_n \xrightarrow{P} \alpha X + \beta Y, \quad \forall \alpha, \beta \in \mathbb{R} \\ \\ X_n \cdot Y_n \xrightarrow{P} X \cdot Y \end{array} \right.$$

6.2.2 Convergence en loi et convergence en probabilité

Théorème 6.2.9

$$X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{\mathcal{L}} X$$

Proposition 6.2.10

$$X_n \xrightarrow{\mathcal{L}} a \in \mathbb{R} \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

Proposition 6.2.11 *Si* $X_n \xrightarrow{\mathcal{L}} X$ *et* $Y_n \xrightarrow{P} a \in \mathbb{R}$ *alors :*

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} X \\ a \end{pmatrix}.$$

En particulier :

(a) $X_n + Y_n \xrightarrow{\mathcal{L}} X + a$,

(b) $X_n \cdot Y_n \xrightarrow{\mathcal{L}} X \cdot a$,

(c) $\dfrac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \dfrac{X}{a}, \ \forall a \in \mathbb{R}^*$.

6.3 Convergence presque sûre

6.3.1 Définition

Définition 6.3.1 *La suite* $(X_n)_{n\in\mathbb{N}}$ *de vecteurs aléatoires* **converge presque sûrement** *vers* X*, et on note* $X_n \xrightarrow{p.s.} X$*, si :*

$$\exists A \in \mathcal{A},\ P(A) = 1\ \ t.q.\ \ X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega),\ \forall \omega \in A.$$

6.3.2 Critères de convergence p.s.

Théorème 6.3.2

$$\sup_{m\geq n} |X_m - X| \xrightarrow[n\to\infty]{P} 0 \quad \Rightarrow \quad X_n \xrightarrow{p.s.} X$$

Proposition 6.3.3

$$\sum_{n=0}^{\infty} P\left(|X_n - X| > \varepsilon\right) < \infty, \quad \forall \varepsilon > 0 \quad \Rightarrow \quad X_n \xrightarrow{p.s.} X$$

6.3.3 Convergence en probabilité et convergence presque sûre

Théorème 6.3.4

$$X_n \xrightarrow{p.s.} X \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

Proposition 6.3.5

$$X_n \xrightarrow{P} X \quad \Rightarrow \quad \exists \left(X_{n_k}\right)_{k \in \mathbb{N}} \subset \left(X_n\right)_{n \in \mathbb{N}} : X_{n_k} \xrightarrow{p.s.} X$$

6.3.4 Loi forte des grands nombres

Théorème 6.3.6 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de vecteurs aléatoires indépendants et de même loi. Soit* μ *l'espérance de cette loi. On a alors*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} \mu.$$

Théorème 6.3.7 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a.r. de* L^2, *indépendantes. Si*

$$E(X_n) \underset{n\to\infty}{\longrightarrow} \mu, \qquad \sum_{n=1}^{\infty} \frac{V(X_n)}{n^2} < \infty,$$

alors

$$\overline{X}_n \xrightarrow{p.s.} \mu.$$

6.4 Convergence dans L^p

Définition 6.4.1 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a.r. dans* L^p *et* X *une v.a.r. dans* L^p*. On dit que* $(X_n)_{n\in\mathbb{N}}$ **converge dans** L^p *vers* X*, et on note* $X_n \xrightarrow{L^p} X$*, si*

$$\|X_n - X\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0.$$

Proposition 6.4.2 $1 \leq p < q$:

$$X_n \xrightarrow{L^q} X \quad \Rightarrow \quad X_n \xrightarrow{L^p} X$$

Corollaire 6.4.3

$$X_n \xrightarrow{L^2} X \quad \Rightarrow \quad X_n \xrightarrow{L^1} X$$

Proposition 6.4.4

$$X_n \xrightarrow{L^1} X \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

Proposition 6.4.5 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a.r. dans* L^2*. Si*

$$E(X_n) \underset{n\to\infty}{\longrightarrow} \mu, \qquad V(X_n) \underset{n\to\infty}{\longrightarrow} 0,$$

alors

$$X_n \xrightarrow{L^2} \mu.$$

Théorème 6.4.6 *(Loi des grands nombres dans* L^2*)*
Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a.r. dans* L^2*, non corrélées et de même loi, d'espérance* μ*. On a alors :*

$$\overline{X}_n \xrightarrow{L^2} \mu.$$

6.5 Résumé

$$
\begin{tikzcd}
L^q \arrow[r, "q\geq p\geq 2"'] & L^p \arrow[r] & L^1 \arrow[dr] & & \\
& & & P \arrow[r] & \mathcal{L} \\
& p.co. \arrow[r] & ps \arrow[ur] & &
\end{tikzcd}
$$

Exercice

On considère n v.a.r. i.i.d. $X_1, \ldots, X_n$ issu d'une loi de probabilité P d'espérance m, de variance σ^2 et de moment centré d'ordre 4 noté $\mu_4 = E\left(X_1 - m\right)^4 < \infty$, tel que $\mu_4 - \sigma^4 > 0$. On définit :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2, \quad S_n'^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2.$$

1. (a) Calculer $E\left(\overline{X}_n\right)$ et $V\left(\overline{X}_n\right)$.

   (b) Quelle est la loi asymptotique de $\sqrt{n}\left(\overline{X}_n - m\right)$ quand $n \to +\infty$?

   (c) Vers quelle valeur converge $\overline{X}_n$ quand $n \to \infty$ (convergence en probabilité ou convergence presque sûre) ?

   (d) Calculer $E\left(S_n^2\right)$, $E\left(S_n'^2\right)$ et $V\left(S_n'^2\right)$.

   (e) Vers quelle limite converge $S_n'^2$? S_n^2 ?

   (f) Appliquer le théorème limite centrale à la suite de v.a. $(X_i - m)^2$ pour en déduire la loi asymptotique de $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - \sigma^2\right)$.

   (g) Peut-on obtenir un résultat de convergence en loi de $\sqrt{n}\left(S_n^2 - \sigma^2\right)$ vers un vecteur aléatoire gaussien, par application du TLC ?

2. (a) Etablir que $cov\left(\overline{X}_n, S_n'^2\right) = \frac{\mu_3}{n}$.

(b) En déduire que si P est une loi symétrique, alors pour tout n les v.a. $\overline{X}_n$ et $S_n'^2$ ne sont pas corrélées linéairement. La réciproque est-elle vraie ?

(c) En général, peut-on conclure que les variables $\overline{X}_n$ et $S_n'^2$ sont asymptotiquement non corrélées ?

3. Applications

(a) On suppose la loi P gaussienne.

i. Montrer que la statistique $T = \dfrac{\sqrt{n}\left(\overline{X}_n - m\right)}{S'_n}$ suit une loi de Student.

ii. Calculer le coefficient de corrélation linéaire ρ entre $\overline{X}_n$ et $T = \dfrac{\sqrt{n}\left(\overline{X}_n - m\right)}{S'_n}$.

iii. Donner la loi exacte de $\overline{X}_n$, de S'^2_n , de $\left(\overline{X}_n, S'^2_n\right)^t$.

(b) On suppose que la loi P appartient à la famille des lois gamma $\gamma(p, \theta)$ de densité :

$$f(x) = \frac{1}{\Gamma(p)} \, e^{-\theta x} \, \theta^{p} \, x^{p-1} \, \mathbb{1}_{\mathbb{R}^{+}}(x),$$

donc $E\left(X_i^r\right) = \dfrac{\Gamma(p+r)}{\theta^r \, \Gamma(p)}$.

i. Quelle est la loi exacte de $\overline{X}_n$?

ii. Quelle est la loi asymptotique de $\overline{X}_n$?

iii. Soit $Y_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i^2$. Vers quelle limite converge Y_n (p.s. ou proba) ?

iv. Etablir la convergence en loi de Y_n et de $\left(\overline{X}_n, Y_n\right)^t$.

7 Compléments

Définition 7.0.1 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $(A_n)_{n\in\mathbb{N}}$ *une suite d'événements de* $\mathcal{A}$*. L'ensemble des* $\omega \in \Omega$ *appartenant à une infinité de* A_n*, soit*

$$\begin{aligned}\overline{\lim} A_n &= \bigcap_{n\geq 0} \bigcup_{k\geq n} A_k = \{\, \omega \in \Omega : \sum_{n\geq 0} \mathbb{1}_{A_n}(\omega) = \infty \,\} \\ &= \{\omega \in \Omega : \exists n,\ \forall k \geq n,\ \omega \in A_k\}\,,\end{aligned}$$

est appelé limite sup des A_n*.*

L'ensemble des $\omega \in \Omega$ *appartenant à tous les* A_n*, sauf peut-être un nombre fini, soit*

$$\begin{aligned}\underline{\lim} A_n &= \bigcup_{n\geq 0} \bigcap_{k\geq n} A_k = \{\, \omega \in \Omega : \sum_{n\geq 0} \mathbb{1}_{\overline{A}_n}(\omega) < \infty \,\} \\ &= \{\omega \in \Omega : \forall n,\ \exists k \geq n,\ \omega \in A_k\}\,,\end{aligned}$$

est appelé limite inf des A_n*.*

Théorème 7.0.2 (Lemme de Borel-Cantelli) *Soit* $(A_n)_{n\in\mathbb{N}}$ *une suite d'événements de* $\mathcal{A}$*.*

1. *Si* $\sum_{n\geq 0} P(A_n) < \infty$ *alors* $P\left(\overline{\text{lim}} A_n\right) = 0$.

2. *Si* $\sum_{n\geq 0} P(A_n) = \infty$ *et si la suite* $(A_n)_{n\in\mathbb{N}}$ *est indépendante, alors* $P\left(\overline{\text{lim}} A_n\right) = 1$.

Exemple 1

Soit $(X_n)_n$ une suite de v.a.r. i.i.d. de loi $\mathcal{B}\left(1, \frac{1}{n}\right)$. Alors

$$X_n \xrightarrow{L^1} 0, \qquad X_n \xrightarrow{P} 0, \qquad X_n \not\xrightarrow{p.s.} 0.$$

Exemple 2

Soit $X \sim \mathcal{N}(0,1)$ et $X_n = -X$, $\forall n \in \mathbb{N}$. Alors

$$X_n \xrightarrow{\mathcal{L}} X, \qquad X_n \not\xrightarrow{P} X.$$

Exemple 3

Soit $(X_n)_n$ une suite de v.a.r. t.q. $P\left(X_n = n\right) = \dfrac{1}{n}$ et $P\left(X_n = 0\right) = 1 - \dfrac{1}{n}$. Alors

$$X_n \xrightarrow{P} 0, \qquad X_n \not\xrightarrow{L^1} 0.$$

Exemple 4

Soit $X_n = n\ \mathbb{1}_{\left(0,\frac{1}{n}\right]}$, $\forall n \in \mathbb{N}$. Alors

$$X_n \xrightarrow{p.s.} 0, \qquad X_n \not\xrightarrow{L^1} 0.$$

Exemple 5 *(Formule de Stirling)*

Soit $(X_n)_n$ une suite de v.a.r. i.i.d. de loi $\mathcal{P}(1)$. Alors

$$S_n = X_1 + \ldots + X_n \sim \mathcal{P}(n) \qquad \Rightarrow \qquad P(S_n = n) = e^{-n}\,\frac{n^n}{n!}.$$

Par le TLC,

$$Z_n = \sqrt{n}\left(\frac{S_n}{n} - 1\right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0,1),$$

donc pour n assez grand

$$\begin{aligned} P(n-1 < S_n \leq n) &= P\left(-\frac{1}{\sqrt{n}} < Z_n \leq 0\right) \simeq P\left(Z \in \left(-\frac{1}{\sqrt{n}}, 0\right)\right) \\ &\simeq \frac{1}{\sqrt{2\pi}} \int_{-\frac{1}{\sqrt{n}}}^{0} e^{-\frac{x^2}{2}}\, dx \simeq \frac{1}{\sqrt{2\pi n}}. \end{aligned}$$

On obtient

$$e^{-n}\ \frac{n^n}{n!} \simeq \frac{1}{\sqrt{2\pi n}} \qquad \Longleftrightarrow \qquad n! \simeq n^n\ e^{-n}\ \sqrt{2\pi n}.$$